Annual Issue 2011

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), Zentralblatt MATH (Germany), 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,
200 College Road, Gallup, NM 87301, USA

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

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

JANUARY 2011 **VOLUME 1**

CONTENTS

Information for Authors and Subscribers

Progress in Physics has been created for 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. 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. We accept brief papers, no larger than 8 typeset journal pages. Short articles are preferable. Large papers can be considered in exceptional cases to the section *Special Reports* intended for such publications in the journal. 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.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

SPECIAL REPORT

Re-Identification of the Many-World Background of Special Relativity as Four-World Background. Part I.

Akindele O. Joseph Adekugbe

Center for The Fundamental Theory, P. O. Box 2575, Akure, Ondo State 340001, Nigeria.
E-mail: adekugbe@alum.mit.edu

The pair of co-existing symmetrical universes, referred to as our (or positive) universe and negative universe, isolated and shown to constitute a two-world background for the special theory of relativity (SR) in previous papers, encompasses another pair of symmetrical universes, referred to as positive time-universe and negative time-universe. The Euclidean 3-spaces (in the context of SR) of the positive time-universe and the negative time-universe constitute the time dimensions of our (or positive) universe and the negative universe respectively, relative to observers in the Euclidean 3-spaces of our universe and the negative universe and the Euclidean 3-spaces of our universe and the negative universe constitute the time dimensions of the positive time-universe and the negative time-universe respectively, relative to observers in the Euclidean 3-spaces of the positive time-universe and the negative time-universe. Thus time is a secondary concept derived from the concept of space according to this paper. The one-dimensional particle or object in time dimension to every three-dimensional particle or object in 3-space in our universe is a three-dimensional particle or object in 3-space in the positive time-universe. Perfect symmetry of natural laws is established among the resulting four universes and two outstanding issues about the new spacetime/intrinsic spacetime geometrical representation of Lorentz transformation/intrinsic Lorentz transformation in the two-world picture, developed in the previous papers, are resolved within the larger four-world picture in this first part of this paper.

1 Origin of time and intrinsic time dimensions

1.1 Orthogonal Euclidean 3-spaces

Let us start with an operational definition of orthogonal Euclidean 3-spaces. Given a three-dimensional Euclidean space (or a Euclidean 3-space) $\boldsymbol{E}^3$ with mutually orthogonal straight line dimensions x^1, x^2 and x^3 and another Euclidean 3-space $\boldsymbol{E}^{03}$ with mutually orthogonal straight line dimensions x^{01}, x^{02} and x^{03}, the Euclidean 3-space $\boldsymbol{E}^{03}$ shall be said to be orthogonal to the Euclidean 3-space $\boldsymbol{E}^3$ if, and only if, each dimension x^{0j} of $\boldsymbol{E}^{03}$; $j = 1, 2, 3$, is orthogonal to every dimension x^i; $i = 1, 2, 3$ of $\boldsymbol{E}^3$. In other words, $\boldsymbol{E}^{03}$ shall be said to be orthogonal to $\boldsymbol{E}^3$ if, and only if, $x^{0j} \perp x^i$; $i, j = 1, 2, 3$, at every point of the Euclidean 6-space generated by the orthogonal Euclidean 3-spaces.

We shall take the Euclidean 3-spaces $\boldsymbol{E}^3$ and $\boldsymbol{E}^{03}$ to be the proper (or classical) Euclidean 3-spaces of classical mechanics (including classical gravity), to be re-denoted by Σ' and $\Sigma^{0\prime}$ respectively for convenience in this paper. The reason for restricting to the proper (or classical) Euclidean 3-spaces is that we shall assume the absence of relativistic gravity while considering the special theory of relativity (SR) on flat spacetime, as shall be discussed further at the end of this paper.

Graphically, let us consider the Euclidean 3-space Σ' with mutually orthogonal straight line dimensions $x^{1\prime}$, $x^{2\prime}$ and $x^{3\prime}$ as a hyper-surface to be represented by a horizontal plane surface and the Euclidean 3-space $\Sigma^{0\prime}$ with mutually orthogonal straight line dimensions $x^{01\prime}$, $x^{02\prime}$ and $x^{03\prime}$ as a hyper-surface to be represented by a vertical plane surface. The union of the two orthogonal proper (or classical) Euclidean 3-spaces yields a compound six-dimensional proper (or classical) Euclidean space with mutually orthogonal dimensions $x^{1\prime}$, $x^{2\prime}$, $x^{3\prime}$, $x^{01\prime}$, $x^{02\prime}$ and $x^{03\prime}$ illustrated in Fig. 1.

As introduced (as *ansatz*) in [1] and as shall be derived formally in the two parts of this paper, the hyper-surface (or proper Euclidean 3-space) Σ' along the horizontal is underlied by an isotropic one-dimensional proper intrinsic space denoted by $\phi\rho'$ (that has no unique orientation in the Euclidean 3-space Σ'). The vertical proper Euclidean 3-space $\Sigma^{0\prime}$ is likewise underlied by an isotropic one-dimensional proper intrinsic space $\phi\rho^{0\prime}$ (that has no unique orientation in the Euclidean 3-space $\Sigma^{0\prime}$). The underlying intrinsic spaces $\phi\rho'$ and $\phi\rho^{0\prime}$ are also shown in Fig. 1.

Inclusion of the proper time dimension ct' along the vertical, normal to the horizontal hyper-surface (or horizontal Euclidean 3-space) Σ' in Fig. 1, yields the flat four-dimensional proper spacetime (Σ', ct') of classical mechanics (CM), (including classical gravitation), of the positive (or our) universe and inclusion of the proper intrinsic time dimension $\phi c\phi t'$ along the vertical, normal to the proper intrinsic space $\phi\rho'$ along the horizontal, yields the flat 2-dimensional proper intrinsic spacetime $(\phi\rho', \phi c\phi t')$ of intrinsic classical mechanics

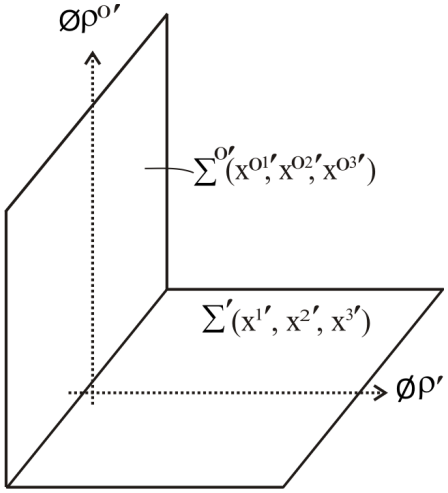


Fig. 1: Co-existing two orthogonal proper Euclidean 3-spaces (considered as hyper-surfaces) and their underlying isotropic one-dimensional proper intrinsic spaces.

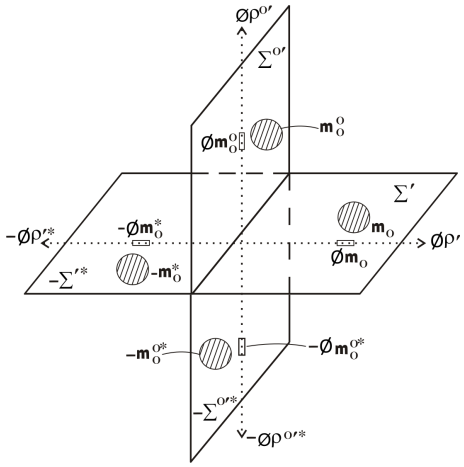


Fig. 2: Co-existing four mutually orthogonal proper Euclidean 3-spaces and their underlying isotropic one-dimensional proper intrinsic spaces, where the rest masses in the proper Euclidean 3-spaces and the one-dimensional intrinsic rest masses in the intrinsic spaces of a quartet of symmetry-partner particles or object are shown.

(ϕCM), (including intrinsic classical gravitation), of our universe. The proper Euclidean 3-space Σ' and its underlying one-dimensional proper intrinsic space $\phi\rho'$ shall sometimes be referred to as our proper (or classical) Euclidean 3-space and our proper (or classical) intrinsic space for brevity.

The vertical proper Euclidean 3-space $\Sigma^{0\prime}$ and its underlying one-dimensional proper intrinsic space $\phi\rho^{0\prime}$ in Fig. 1 are new. They are different from the proper Euclidean 3-space $-\Sigma'^*$ and its underlying proper intrinsic space $-\phi\rho'^*$ of the negative universe isolated in [1] and [2]. The Euclidean 3-space $-\Sigma'^*$ and its underlying proper intrinsic space $-\phi\rho'^*$ of the negative universe are "anti-parallel" to the Euclidean 3-space Σ' and its underlying intrinsic space $\phi\rho'$ of the positive universe, which means that the dimensions $-x^{1\prime*}$, $-x^{2\prime*}$ and $-x^{3\prime*}$ of $-\Sigma'^*$ are inversions in the origin of the dimensions $x^{1\prime}$, $x^{2\prime}$ and $x^{3\prime}$ of Σ'.

There are likewise the proper Euclidean 3-space $-\Sigma^{0\prime*}$ and its underlying proper intrinsic space $-\phi\rho^{0\prime*}$, which are "anti-parallel" to the new proper Euclidean 3-space $\Sigma^{0\prime}$ and its underlying proper intrinsic space $\phi\rho^{0\prime}$ in Fig. 1. Fig. 1 shall be made more complete by adding the negative proper Euclidean 3-spaces $-\Sigma'^*$ and $-\Sigma^{0\prime*}$ and their underlying one-dimensional intrinsic spaces $-\phi\rho'^*$ and $-\phi\rho^{0\prime*}$ to it, yielding Fig. 2.

The proper Euclidean 3-space Σ' with dimensions $x^{1\prime}$, $x^{2\prime}$ and $x^{3\prime}$ and the proper Euclidean 3-space $\Sigma^{0\prime}$ with dimensions $x^{01\prime}$, $x^{02\prime}$ and $x^{03\prime}$ in Fig. 2 are orthogonal Euclidean 3-spaces, which means that $x^{0j\prime} \perp x^{i\prime}$; $i, j = 1, 2, 3$, as defined earlier. The proper Euclidean 3-space $-\Sigma'^*$ with dimensions $-x^{1\prime*}$, $-x^{2\prime*}$ and $-x^{3\prime*}$ and the proper Euclidean 3-space $-\Sigma^{0\prime*}$ with dimensions $-x^{01\prime*}$, $-x^{02\prime*}$ and $-x^{03\prime*}$ are likewise orthogonal Euclidean 3-spaces.

Should the vertical Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime*}$ and their underlying isotropic intrinsic spaces $\phi\rho^{0\prime}$ and $-\phi\rho^{0\prime*}$ exist naturally, then they should belong to a new pair of worlds (or universes), just as the horizontal proper Euclidean 3-space Σ' and $-\Sigma'^*$ and their underlying one-dimensional isotropic proper intrinsic spaces $\phi\rho'$ and $-\phi\rho'^*$ exist naturally and belong to the positive (or our) universe and the negative universe respectively, as found in [1] and [2]. The appropriate names for the new pair of universes with flat four-dimensional proper spacetimes ($\Sigma^{0\prime}$, $ct^{0\prime}$) and ($-\Sigma^{0\prime*}$, $-ct^{0\prime*}$) of classical mechanics (CM) and their underlying flat two-dimensional proper intrinsic spacetimes ($\phi\rho^{0\prime}$, $\phi c\phi t^{0\prime}$) and ($-\phi\rho^{0\prime*}$, $-\phi c\phi t^{0\prime*}$) of intrinsic classical mechanics (ϕCM), where the time dimensions and intrinsic time dimensions have not yet appeared in Fig. 2, shall be derived later in this paper.

As the next step, an assumption shall be made, which shall be justified with further development of this paper, that the four universes encompassed by Fig. 2, with flat four- dimensional proper spacetimes (Σ', ct'), ($-\Sigma'^*, -ct'^*$), ($\Sigma^{0\prime}, ct^{0\prime}$) and ($-\Sigma^{0\prime*}, -ct^{0\prime*}$) and their underlying flat proper intrinsic spacetimes ($\phi\rho', \phi c\phi t'$), ($-\phi\rho'^*, -\phi c\phi t'^*$), ($\phi\rho^{0\prime}, \phi c\phi t^{0\prime}$) and ($-\phi\rho^{0\prime*}, -\phi c\phi t^{0\prime*}$) respectively, where the proper time and proper intrinsic time dimensions have not yet appeared in Fig. 2, exist naturally and exhibit perfect symmetry of state and perfect symmetry of natural laws. Implied by this assumption are the following facts:

1. Corresponding to every given point P in our proper Euclidean 3-space Σ', there are unique symmetry- partner point P^0, P^* and P^{0*} in the proper Euclidean 3-spaces $\Sigma^{0\prime}$, $-\Sigma'^*$ and $-\Sigma^{0\prime*}$ respectively;
2. Corresponding to every particle or object of rest mass m_0 located at a point in our proper Euclidean 3-space Σ', there are identical symmetry-partner particles or ob-

jects of rest masses to be denoted by m_0^0, $-m_0^*$ and $-m_0^{0*}$ located at the symmetry-partner points in $\Sigma^{0\prime}$, $-\Sigma^{\prime *}$ and $-\Sigma^{0\prime *}$ respectively, as illustrated in Fig. 2 already and

3. Corresponding to motion at a speed v of the rest mass m_0 of a particle or object through a point along a direction in our proper Euclidean space $\Sigma^{\prime}$, relative to an observer in $\Sigma^{\prime}$, there are identical symmetry- partner particles or objects of rest masses m_0^0, $-m_0^*$ and $-m_0^{0*}$ in simultaneous motions at equal speed v along identical directions through the symmetry-partner points in the proper Euclidean 3-spaces $\Sigma^{0\prime}$, $-\Sigma^{\prime *}$ and $-\Sigma^{0\prime *}$ respectively, relative to identical symmetry-partner observers in the respective Euclidean 3-spaces.

4. A further requirement of the symmetry of state among the four universes encompassed by Fig. 2 is that the motion at a speed v of a particle along the $X-$ axis, say, of its frame in any one of the four proper Euclidean 3-spaces, (in $\Sigma^{0\prime}$, say), relative to an observer (or frame of reference) in that proper Euclidean 3-space (or universe), is equally valid relative to the symmetry-partner observers in the three other proper Euclidean 3-spaces (or universes). Consequently the simultaneous rotations by equal intrinsic angle $\phi\psi$ of the intrinsic affine space coordinates of the symmetry-partner particles' frames $\phi\tilde{x}^{\prime}$, $\phi\tilde{x}^{0\prime}$, $-\phi\tilde{x}^{\prime *}$ and $-\phi\tilde{x}^{0\prime *}$ relative to the intrinsic affine space coordinates of the symmetry-partner observers' frames $\phi\tilde{x}$, $\phi\tilde{x}^0$, $-\phi\tilde{x}^*$ and $-\phi\tilde{x}^{0*}$ respectively in the context of the intrinsic special theory of relativity (ϕSR), as developed in [1], implied by item 3, are valid relative to every one of the four symmetry-partner observers in the four proper Euclidean 3-spaces (or universes). Thus every one of the four symmetry-partner observers can validly draw the identical relative rotations of affine intrinsic spacetime coordinates of symmetry-partner frames of reference in the four universes encompassed by Fig. 2 with respect to himself and construct ϕSR and consequently SR in his universe with the diagram encompassing the four universes he obtains.

Inherent in item 4 above is the fact that the four universes with flat four-dimensional proper physical (or metric) spacetimes $(\Sigma^{\prime}, ct^{\prime})$, $(\Sigma^{0\prime}, ct^{0\prime})$, $(-\Sigma^{\prime *}, -ct^{\prime *})$ and $(-\Sigma^{0\prime *}, -ct^{0\prime *})$ of classical mechanics (CM) in the universes encompassed by Fig. 2, (where the proper time dimensions have not yet appeared), are stationary dynamically relative to one another at all times. Otherwise the speed v of a particle in a universe (or in a Euclidean 3-space in Fig. 2) relative to an observer in that universe (or in that Euclidean 3-space), will be different relative to the symmetry-partner observer in another universe (or in another Euclidean 3-space), who must obtain the speed of the particle relative to himself as the resultant of the particle's speed v relative to the observer in the particle's universe and the speed V_0 of the particle's universe (or particle's Euclidean 3-space) relative to his universe (or his Euclidean 3-space). The simultaneous identical relative rotations by equal intrinsic angle of intrinsic affine spacetime coordinates of symmetry-partner frames of reference in the four universes, which symmetry of state requires to be valid with respect to every one of the four symmetry-partner observers in the four universes, will therefore be impossible in the situation where some or all the four universes (or Euclidean 3-spaces in Fig. 2) are naturally in motion relative to one another.

Now the proper intrinsic metric space $\phi\rho^{0\prime}$ along the vertical in the first quadrant is naturally rotated at an intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$ relative to the proper intrinsic metric space $\phi\rho^{\prime}$ of the positive (or our) universe along the horizontal in the first quadrant in Fig. 2. The proper intrinsic metric space $-\phi\rho^{\prime *}$ of the negative universe is naturally rotated at intrinsic angle $\phi\psi_0 = \pi$ relative to our proper intrinsic metric space $\phi\rho^{\prime}$ and the proper intrinsic metric space $-\phi\rho^{0\prime *}$ along the vertical in the third quadrant is naturally rotated at intrinsic angle $\phi\psi_0 = \frac{3\pi}{2}$ relative to our proper intrinsic metric space $\phi\rho^{\prime}$ in Fig. 2. The intrinsic angle of natural rotations of the intrinsic metric spaces $\phi\rho^{0\prime}$, $-\phi\rho^{\prime *}$ and $-\phi\rho^{0\prime *}$ relative to $\phi\rho^{\prime}$ has been denoted by $\phi\psi_0$ in order differentiate it from the intrinsic angle of relative rotation of intrinsic affine spacetime coordinates in the context of ϕSR denoted by $\phi\psi$ in [1].

The natural rotations of the one-dimensional proper intrinsic metric spaces $\phi\rho^{0\prime}$, $-\phi\rho^{\prime *}$ and $-\phi\rho^{0\prime *}$ relative to our proper intrinsic metric space $\phi\rho^{\prime}$ at different intrinsic angles $\phi\psi_0$ discussed in the foregoing paragraph, implies that the intrinsic metric spaces $\phi\rho^{0\prime}$, $-\phi\rho^{\prime *}$ and $-\phi\rho^{0\prime *}$ possess different intrinsic speeds, to be denoted by ϕV_0, relative to our intrinsic metric space $\phi\rho^{\prime}$. This is deduced in analogy to the fact that the intrinsic speed ϕv of the intrinsic rest mass ϕm_0 of a particle relative to an observer causes the rotations of the intrinsic affine spacetime coordinates $\phi\tilde{x}^{\prime}$ and $\phi c\phi\tilde{t}^{\prime}$ of the particle's intrinsic frame at equal intrinsic angle $\phi\psi$ relative to the intrinsic affine spacetime coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ respectively of the observe's intrinsic frame in the context of intrinsic special relativity (ϕSR), as developed in [1] and presented graphically in Fig. 8a of that paper.

Indeed the derived relation, $\sin\phi\psi = \phi v/\phi c$, between the intrinsic angle $\phi\psi$ of inclination of the intrinsic affine space coordinate $\phi\tilde{x}^{\prime}$ of the particle's intrinsic frame relative to the intrinsic affine space coordinate $\phi\tilde{x}$ of the observer's intrinsic frame in the context of ϕSR, presented as Eq. (18) of [1], is equally valid between the intrinsic angle $\phi\psi_0$ of natural rotation of a proper intrinsic metric space $\phi\rho^{0\prime}$, say, relative to our proper intrinsic metric space $\phi\rho^{\prime}$ in Fig. 2 and the implied natural intrinsic speed ϕV_0 of $\phi\rho^{0\prime}$ relative to $\phi\rho^{\prime}$. In other words, the following relation obtains between $\phi\psi_0$ and ϕV_0:

$$\sin\phi\psi_0 = \phi V_0/\phi c \tag{1}$$

It follows from (1) that the intrinsic metric space $\phi\rho^{0\prime}$ naturally possesses intrinsic speed $\phi V_0 = \phi c$ relative to our in-

trinsic metric space $\phi\rho'$, which is so since $\phi\rho^{0\prime}$ is naturally inclined at intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$ relative to $\phi\rho'$; the proper intrinsic metric space $-\phi\rho'^*$ of the negative universe naturally possesses zero intrinsic speed ($\phi V_0 = 0$) relative to our proper intrinsic metric space $\phi\rho'$, since $-\phi\rho'^*$ is naturally inclined at intrinsic angle $\phi\psi_0 = \pi$ relative to $\phi\rho'$ and the intrinsic metric space $-\phi\rho^{0\prime*}$ naturally possesses intrinsic speed $\phi V_0 = -\phi c$ relative to our intrinsic metric space $\phi\rho'$, since $-\phi\rho^{0\prime*}$ is naturally inclined at $\phi\psi_0 = \frac{3\pi}{2}$ relative to $\phi\rho'$ in Fig. 2.

On the other hand, $-\phi\rho^{0\prime*}$ possesses positive intrinsic speed $\phi V_0 = \phi c$ relative to $-\phi\rho'^*$, since $-\phi\rho^{0\prime*}$ is naturally inclined at intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$ relative to $-\phi\rho'^*$ and $\phi\rho^{0\prime}$ naturally possesses negative intrinsic speed $\phi V_0 = -\phi c$ relative to $-\phi\rho'^*$, since $\phi\rho^{0\prime}$ is naturally inclined at $\phi\psi_0 = \frac{3\pi}{2}$ relative to $-\phi\rho'^*$ in Fig 2. These facts have been illustrated in Figs. 10a and 10b of [1] for the concurrent open intervals $(-\frac{\pi}{2}, \frac{\pi}{2})$ and $(\frac{\pi}{2}, \frac{3\pi}{2})$ within which the intrinsic angle $\phi\psi$ could take on values with respect to 3-observers in the Euclidean 3-spaces Σ' of the positive universe and $-\Sigma'^*$ of the negative universe.

The natural intrinsic speed $\phi V_0 = \phi c$ of $\phi\rho^{0\prime}$ relative to $\phi\rho'$ will be made manifest in speed $V_0 = c$ of the Euclidean 3-space $\Sigma^{0\prime}$ relative to our Euclidean 3-space Σ'; the natural zero intrinsic speed ($\phi V_0 = 0$) of the intrinsic space $-\phi\rho'^*$ of the negative universe relative to $\phi\rho'$ will be made manifest in natural zero speed ($V_0 = 0$) of the Euclidean 3-space $-\Sigma'^*$ of the negative universe relative to our Euclidean 3-space Σ' and the natural intrinsic speed $\phi V_0 = -\phi c$ of $-\phi\rho^{0\prime*}$ relative to $\phi\rho'$ will be made manifest in natural speed $V_0 = -c$ of the Euclidean 3-space $-\Sigma^{0\prime*}$ relative to our Euclidean 3-space Σ' in Fig. 2. By incorporating the additional information in this and the foregoing two paragraphs into Fig. 2 we have Fig. 3, which is valid with respect to 3-observers in our proper Euclidean 3-spaces Σ', as indicated.

There are important differences between the speeds V_0 of the Euclidean 3-spaces that appear in Fig. 3 and speed v of relative motion of particles and objects that appear in the special theory of relativity (SR). First of all, the speed v of relative motion is a property of the particle or object in relative motion, which exists nowhere in the vast space outside the particle at any given instant. This is so because there is nothing (no action-at-a-distance) in relative motion to transmit the velocity of a particle to positions outside the particle. On the other hand, the natural speed V_0 of a Euclidean 3-space is a property of that Euclidean 3-space, which has the same magnitude at every point of the Euclidean 3-space with or without the presence of a particle or object of any rest mass.

The natural speed V_0 of a Euclidean 3-space is isotropic. This means that it has the same magnitude along every direction of the Euclidean 3-space. This is so because each dimension $x^{0j\prime}$; $j = 1, 2, 3$, of $\Sigma^{0\prime}$ is rotated at equal angle $\psi_0 = \frac{\pi}{2}$ relative to every dimension $x^{i\prime}$; $i = 1, 2, 3$, of Σ', (which implies that each dimension $x^{0j\prime}$ of $\Sigma^{0\prime}$ possesses speed $V_0 = c$ naturally relative to every dimension $x^{i\prime}$ of Σ'), thereby making $\Sigma^{0\prime}$ an orthogonal Euclidean 3-space to Σ'. What should be the natural velocity $\vec{V}_0$ of a Euclidean 3-space has components of equal magnitude V_0 along every direction and at every point in that Euclidean 3-space. On the other hand, the speed v of relative motion of a particle or object is not isotropic because the velocity $\vec{v}$ of relative motion along a direction in a Euclidean 3-space has components of different magnitudes along different directions of that Euclidean 3-space. Only the speed $v = c$ of translation of light (or photon) in space is known to be isotropic.

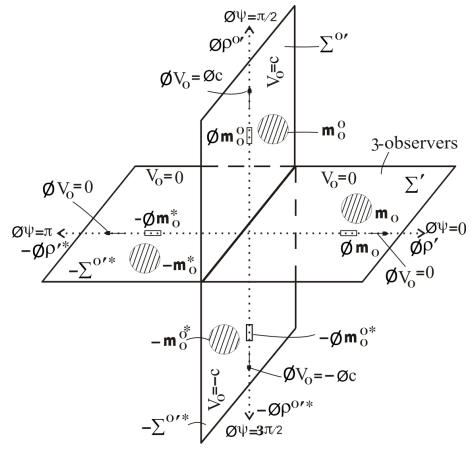


Fig. 3: Co-existing four mutually orthogonal proper Euclidean 3-spaces and their underlying isotropic one-dimensional proper intrinsic spaces, where the speeds V_0 of the Euclidean 3-spaces and the intrinsic speeds ϕV_0 of the intrinsic spaces, relative to 3-observers in our proper Euclidean 3-space (considered as a hyper-surface along the horizontal) in the first quadrant are shown.

Now a material particle or object of any magnitude of rest mass that is located at any point in a Euclidean 3-space acquires the natural speed V_0 of that Euclidean 3-space. Thus the rest mass m_0 of the particle or object located in our proper Euclidean 3-space Σ' possesses the spatially uniform natural zero speed ($V_0 = 0$) of Σ' relative to every particle, object or observer in Σ' in Fig. 3. Likewise the rest mass m_0^0 of a particle or object located at any point in the proper Euclidean 3-space $\Sigma^{0\prime}$ acquires the isotropic and spatially uniform natural speed $V_0 = c$ of $\Sigma^{0\prime}$ relative to every particle, object or observer in our Euclidean 3-space Σ'.

The rest mass $-m_0^*$ located at any point in the proper Euclidean 3-space $-\Sigma'^*$ of the negative universe acquires the spatially uniform natural zero speed ($V_0 = 0$) of $-\Sigma'^*$ relative to all particles, objects and observers in our Euclidean 3-space Σ' and the rest mass $-m_0^{0*}$ of a particle or object located at any point in the proper Euclidean 3-space $-\Sigma^{0\prime*}$ acquires the isotropic and spatially uniform natural speed $V_0 = -c$ of $-\Sigma^{0\prime*}$ relative to all particles, objects and observers in our Euclidean 3-space Σ' in Fig. 3.

However, as deduced earlier, symmetry of state among the four universes whose proper (or classical) Euclidean 3-

spaces appear in Fig. 2 or 3 requires that the four universes must be stationary relative to one another always. Then in order to resolve the paradox ensuing from this and the foregoing two paragraphs namely, all the four universes (or their proper Euclidean 3-spaces in Fig. 2 or 3) are stationary relative to one another always (as required by symmetry of state among the four universes), yet the two universes with flat proper spacetimes $(\Sigma^{0\prime}, ct^{0\prime})$ and $(-\Sigma^{0\prime *}, -ct^{0\prime *})$ naturally possess constant speeds $V_0 = c$ and $V_0 = -c$ respectively relative to the flat spacetime (Σ', ct') of our universe, we must consider the constant speeds $V_0 = c$ and $V_0 = -c$ of the universes with the flat spacetimes $(\Sigma^{0\prime}, ct^{0\prime})$ and $(-\Sigma^{0\prime *}, -ct^{0\prime *})$ respectively relative to our universe (or speeds $V_0 = c$ and $V_0 = -c$ of the Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ respectively relative to our Euclidean 3-space Σ' in Fig. 3) as absolute speeds of non-detectable absolute motion. This way, although the two proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ naturally possess speeds $V_0 = c$ and $V_0 = -c$ respectively relative to our proper Euclidean 3-space Σ', the four proper Euclidean 3-spaces encompassed by Fig. 2 or 3 are stationary dynamically (or translation-wise) relative to one another, as required by symmetry of state among the four universes with the four proper Euclidean 3-spaces in Fig. 2 or Fig. 3.

The fact that the natural speed $V_0 = c$ of the proper Euclidean 3-space $\Sigma^{0\prime}$ relative to our proper Euclidean 3-space Σ' or of the rest mass m_0^0 in $\Sigma^{0\prime}$ relative to the symmetry-partner rest mass m_0 in Σ' is an absolute speed of non-detectable absolute motion is certain. This is so since there is no relative motion involving large speed $V_0 = c$ between the rest mass of a particle in the particle's frame and the rest mass of the particle in the observer's frame, (where m_0^0 is the rest mass of the particle and $\Sigma^{0\prime}$ in which m_0^0 is in motion at speed $V_0 = c$ is the particle's frame, while m_0 is the rest mass of the particle located in the observer's frame Σ' in this analogy, knowing that m_0 and m_0^0 are equal in magnitude).

The observer's frame always contains special-relativistic (or Lorentz transformed) coordinates and parameters in special relativity. On the other hand, non-detectable absolute motion does not alter the proper (or classical) coordinates and parameters, as in the case of the non-detectable natural absolute motion at absolute speed $V_0 = c$ of m_0^0 in $\Sigma^{0\prime}$ relative to m_0 that possesses zero absolute speed ($V_0 = 0$) in Σ' in Fig. 3.

We have derived another important difference between the natural speeds V_0 of the Euclidean 3-spaces that appear in Fig. 3 and the speeds v of relative motions of material particles and objects that appear in SR. This is the fact that the isotropic and spatially uniform speed V_0 of a Euclidean 3-space is an absolute speed of non-detectable absolute motion, while speed v of particles and objects is a speed of detectable relative motion.

Thus the isotropic speed $V_0 = c$ acquired by the rest mass m_0^0 located in the proper Euclidean 3-space $\Sigma^{0\prime}$ relative to its symmetry-partner m_0 and all other particles, objects and observers in our proper Euclidean 3-space Σ' in Fig. 3 is a non-detectable absolute speed. Consequently m_0^0 in $\Sigma^{0\prime}$ does not propagate away at speed $V_0 = c$ in $\Sigma^{0\prime}$ from m_0 in Σ' but remains tied to m_0 in Σ' always, despite its isotropic absolute speed c in $\Sigma^{0\prime}$ relative to m_0 in Σ'. The speed $V_0 = -c$ acquired by the rest mass $-m_0^{0*}$ in the proper Euclidean 3-space $-\Sigma^{0\prime *}$ relative to its symmetry-partner rest mass m_0 and all other particles, objects and observers in our Euclidean 3-space Σ' in Fig. 3 is likewise an absolute speed of non-detectable absolute motion. Consequently $-m_0^{0*}$ in $-\Sigma^{0\prime *}$ does not propagate away at speed $V_0 = -c$ in $-\Sigma^{0\prime *}$ from m_0 in Σ' but remains tied to m_0 in Σ' always, despite the absolute speed $V_0 = -c$ of $-m_0^{0*}$ in $-\Sigma^{0\prime *}$ relative to m_0 in Σ'.

On the other hand, the rest mass $-m_0^{0*}$ in $-\Sigma^{0\prime *}$ possesses positive absolute speed $V_0 = c$ and rest mass m_0^0 in $\Sigma^{0\prime}$ possesses negative absolute speed $V_0 = -c$ with respect to the symmetry-partner rest mass $-m_0^*$ and all other particles, objects and observers in $-\Sigma'^*$. This is so since the proper intrinsic space $-\phi\rho^{0\prime *}$ underlying $-\Sigma^{0\prime *}$ is naturally rotated by intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$ relative to the proper intrinsic space $-\phi\rho'^*$ underlying $-\Sigma'^*$ and $\phi\rho^{0\prime}$ underlying $\Sigma^{0\prime}$ is naturally rotated by intrinsic angle $\phi\psi_0 = \frac{3\pi}{2}$ relative to $-\phi\rho'^*$, as mentioned earlier. Consequently $-\phi\rho^{0\prime *}$ naturally possesses absolute intrinsic speed $\phi V_0 = \phi c$ relative to $-\phi\rho'^*$ and $\phi\rho^{0\prime}$ naturally possesses absolute intrinsic speed $\phi V_0 = -\phi c$ relative to $-\phi\rho'^*$. These are then made manifest outwardly as the absolute speed $V_0 = c$ of the Euclidean 3-space $-\Sigma^{0\prime *}$ and absolute speed $V_0 = -c$ of the Euclidean 3-space $\Sigma^{0\prime}$ respectively relative to the Euclidean 3-space $-\Sigma'^*$ of the negative universe.

Let the quartet of symmetry-partner particles or objects of rest masses m_0 in Σ', m_0^0 in $\Sigma^{0\prime}$, $-m_0^*$ in $-\Sigma'^*$ and $-m_0^{0*}$ in $-\Sigma^{0\prime *}$ be located at initial symmetry-partner positions P_i, P_i^0, P_i^* and P_i^{0*} respectively in their respective Euclidean 3-spaces. Then let the particle or object of rest mass m_0 in Σ' be in motion at constant speed v along the $\tilde{x}'-$axis of its frame in our proper Euclidean 3-space Σ' relative to a 3-observer in Σ'. The symmetry-partner particle or object of rest mass m_0^0 in $\Sigma^{0\prime}$ will be in simultaneous motion at equal speed v along the $\tilde{x}^{0\prime}-$axis of its frame in $\Sigma^{0\prime}$ relative to the symmetry-partner observer in $\Sigma^{0\prime}$; the symmetry-partner particle or object of rest mass $-m_0^*$ in $-\Sigma'^*$ will be in simultaneous motion at equal speed v along the $-\tilde{x}'^*-$axis of its frame in $-\Sigma'^*$ relative to the symmetry-partner 3-observer in $-\Sigma'^*$ and the symmetry-partner particle or object of rest mass $-m_0^{0*}$ in $-\Sigma^{0\prime *}$ will be in simultaneous motion at equal speed v along the $-\tilde{x}^{0\prime *}-$axis of its frame in $-\Sigma^{0\prime *}$ relative to the symmetry-partner 3-observer in $-\Sigma^{0\prime *}$.

Thus after a period of time of commencement of motion, the quartet of symmetry-partner particles or objects have covered equal distances along the identical directions of motion in their respective proper Euclidean 3-spaces to arrive at new symmetry-partner positions P, P^0, P^* and P^{0*} in their respective proper Euclidean 3-spaces. This is possible because the four Euclidean 3-spaces are stationary relative to one an-

other always. The quartet of symmetry-partner particles or objects are consequently located at symmetry-partner positions in their respective proper Euclidean 3-spaces always, even when they are in motion relative to symmetry-partner frames of reference in their respective proper Euclidean 3-spaces.

The speed $V_0 = c$ of the proper Euclidean 3-space $\Sigma^{0\prime}$ relative to our proper Euclidean 3-space Σ' is the outward manifestation of the intrinsic speed $\phi V_0 = \phi c$ of the intrinsic metric space $\phi\rho^{0\prime}$ underlying $\Sigma^{0\prime}$ relative to our intrinsic metric space $\phi\rho'$ and relative to our Euclidean 3-space Σ' in Fig. 3. Then since $V_0 = c$ is absolute and is the same at every point of the Euclidean 3-space $\Sigma^{0\prime}$, the intrinsic speed $\phi V_0 = \phi c$ of $\phi\rho^{0\prime}$ relative to $\phi\rho'$ and Σ' is absolute and is the same at every point along the length of $\phi\rho^{0\prime}$. The intrinsic speed $\phi V_0 = -\phi c$ of the intrinsic metric space $-\phi\rho^{0\prime *}$ relative to our intrinsic metric space $\phi\rho'$ and relative to our Euclidean 3-space Σ' is likewise an absolute intrinsic speed and is the same at every point along the length of $-\phi\rho^{0\prime *}$. The zero intrinsic speed ($\phi V_0 = 0$) of the intrinsic metric space $-\phi\rho'^*$ of the negative universe relative to our intrinsic metric space $\phi\rho'$ and relative to our Euclidean 3-space Σ' is the same along the length of $-\phi\rho'^*$.

It follows from the foregoing paragraph that although the proper intrinsic metric spaces $\phi\rho^{0\prime}$ and $-\phi\rho^{0\prime *}$ along the vertical possess intrinsic speeds $\phi V_0 = \phi c$ and $\phi V_0 = -\phi c$ respectively, relative to our proper intrinsic metric space $\phi\rho'$ and relative to our Euclidean 3-space Σ', the four intrinsic metric spaces $\phi\rho'$, $\phi\rho^{0\prime}$, $-\phi\rho'^*$ and $-\phi\rho^{0\prime *}$ in Fig. 3 are stationery relative to one another always, since the intrinsic speeds $\phi V_0 = \phi c$ of $\phi\rho^{0\prime}$ and $\phi V_0 = -\phi c$ of $-\phi\rho'^*$ relative to our intrinsic metric space $\phi\rho'$ and our Euclidean 3-space Σ' are absolute intrinsic speeds, which are not made manifest in actual intrinsic motion.

Likewise, although the intrinsic rest mass ϕm_0^0 in $\phi\rho^{0\prime}$ acquires the intrinsic speed $\phi V_0 = \phi c$ of $\phi\rho^{0\prime}$, it is not in intrinsic motion at the intrinsic speed ϕc along $\phi\rho^{0\prime}$, since the intrinsic speed ϕc it acquires is an absolute intrinsic speed. The absolute intrinsic speed $\phi V_0 = -\phi c$ acquired by the intrinsic rest mass $-\phi m_0^{0*}$ in $-\phi\rho^{0\prime *}$ is likewise not made manifest in actual intrinsic motion of $-\phi m_0^{0*}$ along $-\phi\rho^{0\prime *}$. Consequently the quartet of intrinsic rest masses $\phi m_0, \phi m_0^0, -\phi m_0^*$ and $-\phi m_0^{0*}$ of symmetry-partner particles or objects in the quartet of intrinsic metric spaces $\phi\rho'$, $\phi\rho^{0\prime}$, $-\phi\rho'^*$ and $-\phi\rho^{0\prime *}$, are located at symmetry-partner points in their respective intrinsic spaces always, even when they are in intrinsic motions relative to symmetry-partner frames of reference in their respective Euclidean 3-spaces.

There is a complementary diagram to Fig. 3, which is valid with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ along the vertical, which must also be drawn along with Fig. 3. Now given the quartet of the proper physical (or metric) Euclidean 3-spaces and their underlying one-dimensional intrinsic metric spaces in Fig. 2, then Fig. 3 with the absolute speeds V_0 of the proper Euclidean 3-spaces and absolute intrinsic speed ϕV_0 of the proper intrinsic spaces assigned with respect to 3-observers in the proper Euclidean 3-space Σ' of the positive (or our) universe, ensues automatically.

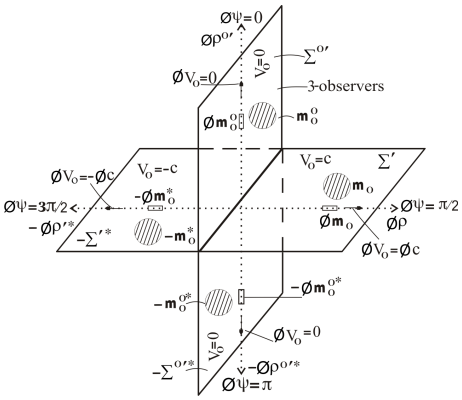


Fig. 4: Co-existing four mutually orthogonal proper Euclidean 3-spaces and their underlying isotropic one-dimensional proper intrinsic metric spaces, where the speeds V_0 of the Euclidean 3-spaces and the intrinsic speeds ϕV_0 of the intrinsic spaces, relative to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ (considered as a hyper-surface) along the vertical in the first quadrant are shown.

On the other hand, the proper physical Euclidean 3-space $\Sigma^{0\prime}$ along the vertical in Fig. 2 possesses zero absolute speed ($V_0 = 0$) at every point of it and its underlying one- dimensional intrinsic space $\phi\rho^{0\prime}$ possesses zero absolute intrinsic speed ($\phi V_0 = 0$) at every point along its length with respect to 3-observers in $\Sigma^{0\prime}$. This is so since $\phi\rho^{0\prime}$ must be considered as rotated by zero intrinsic angle ($\phi\psi_0 = 0$) relative to itself (or relative to the vertical) when the observers of interest are the 3-observers in $\Sigma^{0\prime}$. Then letting $\phi\psi_0 = 0$ in (1) gives zero absolute intrinsic speed ($\phi V_0 = 0$) at every point along $\phi\rho^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$. The physical Euclidean 3-space $\Sigma^{0\prime}$ then possesses zero absolute speed ($V_0 = 0$) at every point of it as the outward manifestation of $\phi V_0 = 0$ at every point along $\phi\rho^{0\prime}$, with respect to 3-observers in $\Sigma^{0\prime}$. It then follows that Fig. 3 with respect to 3-observers in our Euclidean 3-space Σ' corresponds to Fig. 4 with respect to 3-observers in the Euclidean 3-space $\Sigma^{0\prime}$.

It is mandatory to consider the intrinsic metric space $\phi\rho'$ of the positive (or our) universe along the horizontal in the first quadrant as naturally rotated clockwise by a positive intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$; the intrinsic metric space $-\phi\rho^{0\prime *}$ along the vertical in the fourth quadrant as naturally rotated clockwise by a positive intrinsic angle $\phi\psi_0 = \pi$ and the intrinsic metric space $-\phi\rho'^*$ of the negative universe along the horizontal in the third quadrant as naturally rotated clockwise by a positive intrinsic angle $\phi\psi_0 = \frac{3\pi}{2}$ relative to $\phi\rho^{0\prime}$ along the vertical in the first quadrant or with respect to 3-observers in the Euclidean 3-space $\Sigma^{0\prime}$, as indicated in Fig. 4. This way, the positive signs of our proper intrinsic space $\phi\rho'$ and of the

dimensions $x^{1\prime}$, $x^{2\prime}$ and $x^{3\prime}$ of our proper Euclidean 3-space Σ', as well as the positive signs of parameters in Σ' in our (or positive) universe in Fig. 3 are preserved in Fig. 4. The negative signs of $-\phi\rho'^{*}$, $-\Sigma'^{*}$ and of parameters in $-\Sigma'^{*}$ in the negative universe in Fig. 3 are also preserved in Fig. 4, by virtue of the clockwise sense of rotation by positive intrinsic angle $\phi\psi_0$ of $-\phi\rho^{0\prime *}$ and $-\phi\rho'^{*}$ relative to $\phi\rho^{0\prime}$ or with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 4.

If the clockwise rotations of $\phi\rho'$, $-\phi\rho^{0\prime *}$ and $-\phi\rho'^{*}$ relative to $\phi\rho^{0\prime}$ or with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 4, have been considered as rotation by negative intrinsic angles $\phi\psi_0 = -\frac{\pi}{2}, \phi\psi_0 = -\pi$ and $\phi\psi_0 = -\frac{3\pi}{2}$ respectively, then the positive sign of $\phi\rho', \Sigma'$ and of parameters in Σ' of the positive (or our) universe in Fig. 3 would have become negative sign in Fig. 4 and the negative sign of $-\phi\rho'^{*}$ and $-\Sigma'^{*}$ and of parameters in $-\Sigma'^{*}$ of the negative universe in Fig. 3 would have become positive sign in Fig. 4. That is, the positions of the positive and negative universes in Fig. 3 would have been interchanged in Fig. 4, which must not be.

We have thus been led to an important conclusion that natural rotations of intrinsic metric spaces by positive absolute intrinsic angle $\phi\psi_0$ (and consequently the relative rotations of intrinsic affine space coordinates in the context of intrinsic special relativity (ϕSR) by positive relative intrinsic angles, $\phi\psi$), are clockwise rotations with respect to 3-observers in the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ along the vertical (in Fig. 4). Whereas rotation of intrinsic metric spaces (and intrinsic affine space coordinates in the context of ϕSR) by positive intrinsic angles are anti-clockwise rotations with respect to 3-observers in the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ of the positive and negative universes along the horizontal in Fig. 3.

The origin of the natural isotropic absolute speeds V_0 of every point of the proper Euclidean 3-spaces and of the natural absolute intrinsic speeds ϕV_0 of every point along the lengths of the one-dimensional proper intrinsic spaces with respect to the indicated observers in Fig. 3 and Fig. 4, cannot be exposed in this paper. It must be regarded as an outstanding issue to be resolved elsewhere with further development. Nevertheless, a preemptive statement about their origin is appropriate at this point: They are the outward manifestations in the proper physical Euclidean 3-spaces and proper intrinsic spaces of the absolute speeds with respect to the indicated observers, of homogeneous and isotropic absolute spaces (distinguished co-moving coordinate systems) that underlie the proper physical Euclidean 3-spaces and their underlying proper intrinsic spaces in nature, which have not yet appeared in Figs. 3 and Fig. 4.

Leibnitz pointed out that Newtonian mechanics prescribes a distinguished coordinate system (the Newtonian absolute space) in which it is valid [3, see p. 2]. Albert Einstein said, "Newton might no less well have called his absolute space ether..." [4] and argued that the proper (or classical) physical Euclidean 3-space (of Newtonian mechanics) will be impossible without such ether. He also pointed out the existence of ether of general relativity as a necessary requirement for the possibility of that theory, just as the existence of luminiferous ether was postulated to support the propagation of electromagnetic waves. Every dynamical or gravitational law (including Newtonian mechanics) requires (or has) an ether. It is the non-detectable absolute speeds of the ethers of classical mechanics (known to Newton as absolute spaces), which underlie the proper physical Euclidean 3-spaces with respect to the indicated observers in Fig. 3 and 4, that are made manifest in the non-detectable absolute speeds V_0 of the proper Euclidean 3-spaces with respect to the indicated observers in those figures. However this a matter to be formally derived elsewhere, as mentioned above.

1.2 Geometrical contraction of the vertical Euclidean 3-spaces to one-dimensional spaces relative to 3-observers in the horizontal Euclidean 3-spaces and conversely

Let us consider the $x'y'-$plane of our proper Euclidean 3-space Σ' in Fig. 3: Corresponding to the $x'y'-$plane of Σ' is the $x^{0\prime}y^{0\prime}-$plane of the Euclidean 3-space $\Sigma^{0\prime}$. However since Σ' and $\Sigma^{0\prime}$ are orthogonal Euclidean 3-spaces, following the operational definition of orthogonal Euclidean 3-spaces at the beginning of the preceding sub-section, the dimensions $x^{0\prime}$ and $y^{0\prime}$ of the $x^{0\prime}y^{0\prime}-$plane of $\Sigma^{0\prime}$ are both perpendicular to each of the dimensions x' and y' of Σ'. Hence $x^{0\prime}$ and $y^{0\prime}$ are effectively parallel dimensions normal to the $x'y'-$ plane of Σ' with respect to 3-observers in Σ'. Symbolically:

$$x^{0\prime} \perp x' \text{ and } y^{0\prime} \perp x';\ x^{0\prime} \perp y' \text{ and } y^{0\prime} \perp y' \Rightarrow x^{0\prime} \| y^{0\prime} \quad (*)$$

Likewise, corresponding to the $x'z'-$plane of Σ' is the $x^{0\prime}z^{0\prime}-$plane of $\Sigma^{0\prime}$. Again the dimensions $x^{0\prime}$ and $z^{0\prime}$ of the $x^{0\prime}z^{0\prime}-$plane of $\Sigma^{0\prime}$ are both perpendicular to each of the dimensions x' and z' of the $x'z'-$ plane of Σ'. Hence $x^{0\prime}$ and $y^{0\prime}$ are effectively parallel dimensions normal to the $x'z'-$plane of Σ' with respect to 3-observers in Σ'. Symbolically:

$$x^{0\prime} \perp x' \text{ and } z^{0\prime} \perp x';\ x^{0\prime} \perp z' \text{ and } z^{0\prime} \perp z' \Rightarrow x^{0\prime} \| z^{0\prime} \quad (**)$$

Finally, corresponding to the $y'z'-$plane of Σ' is the $y^{0\prime}z^{0\prime}-$plane of $\Sigma^{0\prime}$. Again the dimensions $y^{0\prime}$ and $z^{0\prime}$ of the $y^{0\prime}z^{0\prime}-$plane of $\Sigma^{0\prime}$ are both perpendicular to each of the dimensions y' and z' of the $y'z'-$ plane of Σ'. Hence $y^{0\prime}$ and $z^{0\prime}$ are effectively parallel dimensions normal to the $y'z'-$plane of Σ' with respect to 3-observers in Σ'. Symbolically:

$$y^{0\prime} \perp y' \text{ and } z^{0\prime} \perp y';\ y^{0\prime} \perp z' \text{ and } z^{0\prime} \perp z' \Rightarrow y^{0\prime} \| z^{0\prime} \quad (***)$$

Indeed $x^{0\prime} \| y^{0\prime}$ and $x^{0\prime} \| z^{0\prime}$ in $(*)$ and $(**)$ already implies $y^{0\prime} \| z^{0\prime}$ in $(***)$.

The combination of $(*)$, $(**)$ and $(***)$ give $x^{0\prime} \| y^{0\prime} \| z^{0\prime}$ with respect to 3-observers in Σ', which says that the mutually perpendicular dimensions $x^{0\prime}$, $y^{0\prime}$ and $z^{0\prime}$ of $\Sigma^{0\prime}$ with

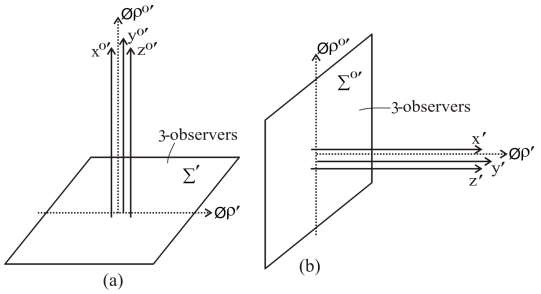


Fig. 5: Given the two orthogonal proper Euclidean 3-spaces $\Sigma^{0\prime}$ and Σ' of Fig. 1 then, (a) the mutually perpendicular dimensions of the proper Euclidean 3-space $\Sigma^{0\prime}$ with respect to 3-observers in it, are naturally "bundle" into parallel dimensions relative to 3-observers in our proper Euclidean 3-space Σ' and (b) the mutually perpendicular dimensions of our proper Euclidean 3-space Σ' with respect to 3-observer in it are naturally "bundled" into parallel dimensions relative to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$.

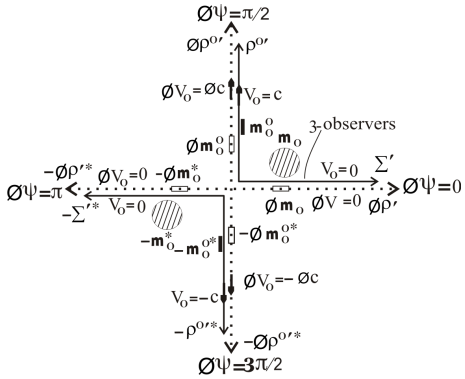


Fig. 6: (a) The proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ along the vertical in Fig. 3, are naturally contracted to one-dimensional proper spaces $\rho^{0\prime}$ and $-\rho^{0\prime *}$ respectively relative to 3-observers in the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ along the horizontal.

respect to 3-observers in $\Sigma^{0\prime}$ are effectively parallel dimensions with respect to 3-observers in our Euclidean 3-space Σ'. In other words, the dimensions $x^{0\prime}$, $y^{0\prime}$ and $z^{0\prime}$ of $\Sigma^{0\prime}$ effectively form a "bundle", which is perpendicular to each of the dimensions x', y' and z' of Σ' with respect to 3-observers in Σ' in Fig. 3. The "bundle" must lie along a fourth dimension with respect to 3-observers in Σ' consequently, as illustrated in Fig. 5a, where the proper Euclidean 3-space Σ' is considered as a hyper-surface represented by a horizontal plane surface.

Conversely, the mutually perpendicular dimensions x', y' and z' of our Euclidean 3-space Σ' with respect to 3-observers in Σ' are effectively parallel dimensions with respect to 3-observers in the Euclidean 3-space $\Sigma^{0\prime}$ in Fig. 4. In other words, the dimensions x', y' and z' of Σ' effectively form a "bundle", which is perpendicular to each of the dimensions $x^{0\prime}$, $y^{0\prime}$ and $z^{0\prime}$ of $\Sigma^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 4. The "bundle" of x', y' and z' must lie along a fourth dimension with respect to 3-observers in $\Sigma^{0\prime}$ consequently, as illustrated in Fig. 5b, where the proper Euclidean 3-space $\Sigma^{0\prime}$ is considered as a hyper-surface represented by a vertical plane surface.

The three dimensions $x^{0\prime}$, $y^{0\prime}$ and $z^{0\prime}$ that are shown as separated parallel dimensions, thereby constituting a "bundle" along the vertical with respect to 3-observers in Σ' in Fig. 5a, are not actually separated. Rather they lie along the singular fourth dimension, thereby constituting a one-dimensional space to be denoted by $\rho^{0\prime}$ with respect to 3-observers in Σ' in Fig. 5a. Likewise the "bundle" of parallel dimensions x', y' and z' effectively constitutes a one-dimensional space to be denoted by ρ' with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 5b. Thus Fig. 5a shall be replaced with the fuller diagram of Fig. 6a, which is valid with respect to 3-observers in the Euclidean 3-space Σ', while Fig. 5b shall be replaced with the fuller diagram of Fig. 6b, which is valid with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$.

Representation of the Euclidean spaces Σ', $-\Sigma'^{*}$, $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ by plane surfaces in the previous diagrams in this paper has temporarily been changed to lines in Figs. 6a and 6b for convenience. The three-dimensional rest masses m_0 and $-m_0^*$ in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ and m_0^0 and $-m_0^{0*}$ in $\Sigma^{0\prime}$and $-\Sigma^{0\prime *}$ have been represented by circles to remind us of their three-dimensionality, while the one-dimensional intrinsic rest masses in the one-dimensional intrinsic spaces $\phi\rho^{0\prime}$, $-\phi\rho^{0\prime *}$, $\phi\rho'$ and $-\phi\rho'^{*}$ and the one-dimensional rest masses in the one-dimensional spaces $\rho^{0\prime}$, $-\rho^{0\prime *}$, ρ' and $-\rho'^{*}$ have been represented by short line segments in Figs. 6a and 6b.

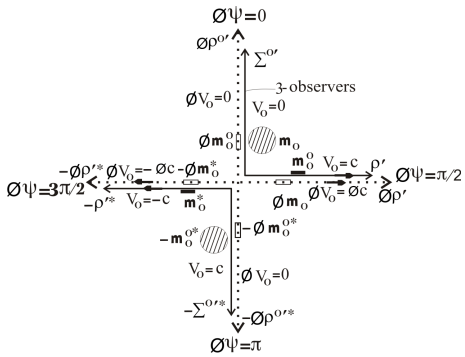


Fig. 6: (b) The proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ along the horizontal in Fig. 4, are naturally contracted to one-dimensional proper spaces ρ' and $-\rho'^{*}$ respectively relative to 3-observers in the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ along the vertical.

Fig. 3 naturally simplifies as Fig. 6a with respect to 3-observers in the proper Euclidean 3-space Σ' of the positive (or our) universe, while Fig. 4 naturally simplifies as Fig. 6b with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ along the vertical. The vertical Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ in Fig. 3 have been geometrically contracted to one-dimensional proper spaces $\rho^{0\prime}$ and $-\rho^{0\prime *}$ respectively with respect to 3-observers in the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ of the positive and negative universes and the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ of the positive and negative universes along the horizontal in Fig. 4, have been geometrically contracted to one-dimensional proper spaces ρ' and $-\rho'^{*}$ respectively with respect to 3-observers in the vertical proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ in Fig. 6b, as actualization of the topic of this sub-section.

The isotropic absolute speed $V_0 = c$ of every point of the Euclidean 3-space $\Sigma^{0\prime}$ with respect to 3-observers in the Euclidean 3-space Σ' in Fig. 3 is now absolute speed $V_0 = c$ of every point along the one-dimensional space $\rho^{0\prime}$ with respect to 3-observers in Σ' in Fig. 6a. The isotropic absolute speed $V_0 = -c$ of every point of the Euclidean 3-space $-\Sigma^{0\prime *}$ with respect to 3-observers in Σ' in Fig. 3 is likewise absolute speed $V_0 = -c$ of every point along the one-dimensional space $-\rho^{0\prime *}$ with respect to 3-observers in Σ' in Fig. 6a.

Just as the absolute speed $V_0 = c$ of every point along $\rho^{0\prime}$ and the absolute intrinsic speed $\phi V_0 = \phi c$ of every point along the intrinsic space $\phi\rho^{0\prime}$ with respect to 3-observers in Σ' in Fig. 6a are isotropic, that is, without unique orientation in the Euclidean 3-space $\Sigma^{0\prime}$ that contracts to $\rho^{0\prime}$, with respect to 3-observers in Σ' and $-\Sigma'^{*}$, so are the one-dimensional space $\rho^{0\prime}$ and the one-dimensional intrinsic space $\phi\rho^{0\prime}$ isotropic dimension and isotropic intrinsic dimension respectively with no unique orientation in the Euclidean 3-space $\Sigma^{0\prime}$, with respect to 3-observers in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$. The one-dimensional space $-\rho^{0\prime *}$ and one-dimensional intrinsic space $-\phi\rho^{0\prime *}$ are likewise isotropic dimension and isotropic intrinsic dimension respectively with no unique orientation in the Euclidean 3-space $-\Sigma^{0\prime *}$ with respect to 3-observers in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ in Fig. 6a.

The isotropic absolute speed $V_0 = c$ of every point of the Euclidean 3-space Σ' and the isotropic absolute speed $V_0 = -c$ of every point of the Euclidean 3-space $-\Sigma'^{*}$ with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 4 are now absolute speed $V_0 = c$ of every point along the one-dimensional space ρ' and absolute speed $V_0 = -c$ of every point along the one-dimensional space $-\rho'^{*}$ with respect to 3-observers in $\Sigma^{0\prime}$ Fig. 6b. Again the one-dimensional metric spaces ρ' and $-\rho'^{*}$ and the one-dimensional intrinsic metric spaces $\phi\rho'$ and $-\phi\rho'^{*}$ are isotropic dimensions and isotropic intrinsic dimensions respectively with no unique orientations in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ that contract to ρ' and $-\rho'^{*}$ respectively, with respect to 3-observers in the vertical Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ in Fig. 6b.

1.3 The vertical proper Euclidean 3-spaces as proper time dimensions relative to 3-observers in the horizontal proper Euclidean 3-spaces and conversely

Figs. 6a and 6b are intermediate diagrams. It shall be shown finally in this section that the one-dimensional proper spaces $\rho^{0\prime}$ and $-\rho^{0\prime *}$ in Fig. 6a naturally transform into the proper time dimensions ct' and $-ct'^{*}$ respectively and their underlying one-dimensional proper intrinsic spaces $\phi\rho^{0\prime}$ and $-\phi\rho^{0\prime *}$ naturally transform into the proper intrinsic time dimensions $\phi c\phi t'$ and $-\phi c\phi t'^{*}$ respectively with respect to 3-observers in the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ in that figure. It shall also be shown that the one-dimensional proper spaces ρ' and $-\rho'^{*}$ in Fig. 6b naturally transform into the proper time dimensions $ct^{0\prime}$ and $-ct^{0\prime *}$ respectively and their underlying proper intrinsic spaces $\phi\rho'$ and $-\phi\rho'^{*}$ naturally transform into proper intrinsic time dimensions $\phi c\phi t^{0\prime}$ and $-\phi c\phi t^{0\prime *}$ respectively with respect to 3-observers in the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ in that figure.

Now let us re-present the generalized forms of the intrinsic Lorentz transformations and its inverse derived and presented as systems (44) and (45) of [1] respectively as follows

$$\left.\begin{aligned} \phi c\,\phi\tilde{t}' &= \sec\phi\psi(\phi c\phi\tilde{t} - \phi\tilde{x}\sin\phi\psi);\\ &\quad (\text{w.r.t. } 1-\text{observer in } c\tilde{t});\\ \phi\tilde{x}' &= \sec\phi\psi(\phi\tilde{x} - \phi c\phi\tilde{t}\sin\phi\psi);\\ &\quad (\text{w.r.t. } 3-\text{observer in } \tilde{\Sigma}) \end{aligned}\right\} \tag{2}$$

and

$$\left.\begin{aligned} \phi c\phi\tilde{t} &= \sec\phi\psi(\phi c\phi\tilde{t}' + \phi\tilde{x}'\sin\phi\psi);\\ &\quad (\text{w.r.t. } 3-\text{observer in } \tilde{\Sigma}');\\ \phi\tilde{x} &= \sec\phi\psi(\phi\tilde{x}' + \phi c\phi\tilde{t}'\sin\phi\psi);\\ &\quad (\text{w.r.t. } 1-\text{observer in } c\tilde{t}') \end{aligned}\right\}. \tag{3}$$

As explained in [1], systems (2) and (3) can be applied for all values of the intrinsic angle $\phi\psi$ in the first cycle, $0 \le \phi\psi \le 2\pi$, except that $\phi\psi = \frac{\pi}{2}$ and $\phi\psi = \frac{3\pi}{2}$ must be avoided.

One observes from system (2) that the pure intrinsic affine time coordinate $\phi c\phi\tilde{t}'$ of the primed (or particle's) intrinsic frame with respect to an observer at rest relative to the particle's frame, transforms into an admixture of intrinsic affine time and intrinsic affine space coordinates of the unprimed (or observer's) intrinsic frame. The pure intrinsic affine space coordinate $\phi\tilde{x}'$ of the primed (or particle's) frame likewise transforms into an admixture of intrinsic affine space and intrinsic affine time coordinates of the unprimed (or particle's) intrinsic frame, when the particle's frame is in motion relative to the observer's frame. The inverses of these observations obtain from system (3), which is the inverse to system (2).

The observations made from system (2) and system (3) described in the foregoing paragraph make the concept of intrinsic affine spacetime induction relevant in relative intrinsic motion of two intrinsic spacetime frames of reference. In order to make this more explicit, let us re-write systems (2) and

(3) respectively as follows

$$\left.\begin{array}{rcl} \phi c\phi\tilde{t}' & = & \sec\phi\psi(\phi c\phi\tilde{t}+\phi c\phi\tilde{t}_i); \\ & & (\text{w.r.t. } 1-\text{observer in } c\tilde{t}); \\ \phi\tilde{x}' & = & \sec\phi\psi(\phi\tilde{x}+\phi\tilde{x}_i); \\ & & (\text{w.r.t. } 3-\text{observer in } \tilde{\Sigma}) \end{array}\right\} \quad (4)$$

and

$$\left.\begin{array}{rcl} \phi c\phi\tilde{t} & = & \sec\phi\psi(\phi c\phi\tilde{t}'+\phi c\phi\tilde{t}'_i); \\ & & (\text{w.r.t. } 3-\text{observer in } \tilde{\Sigma}'); \\ \phi\tilde{x} & = & \sec\phi\psi(\phi\tilde{x}'+\phi\tilde{x}'_i); \\ & & (\text{w.r.t. } 1-\text{observer in } c\tilde{t}') \end{array}\right\}. \quad (5)$$

A comparison of systems (4) and (2) gives the relations for the induced unprimed intrinsic affine spacetime coordinates $\phi c\phi\tilde{t}_i$ and $\phi\tilde{x}_i$ as follows

$$\phi c\phi\tilde{t}_i = \phi\tilde{x}\sin(-\phi\psi) = -\phi\tilde{x}\sin\phi\psi = -\frac{\phi v}{\phi c}\phi\tilde{x}; \quad (6)$$

w.r.t. $1-$ observer in $c\tilde{t}$ and

$$\phi\tilde{x}_i = \phi c\phi\tilde{t}\sin(-\phi\psi) = -\phi c\phi\tilde{t}\sin\phi\psi = -\frac{\phi v}{\phi c}\phi c\phi\tilde{t} = -\phi v\phi\tilde{t}; \quad (7)$$

w.r.t $3-$ observer in $\tilde{\Sigma}$.

Diagrammatically, the induced unprimed intrinsic affine time coordinate, $\phi c\phi\tilde{t}_i = \phi\tilde{x}\sin(-\phi\psi)$ in (6), appears in the fourth quadrant in Fig. 9a of [1] as $\phi\tilde{x}^*\sin(-\phi\psi)$ and the induced unprimed intrinsic affine space coordinate, $\phi\tilde{x}_i = \phi c\phi\tilde{t}\sin(-\phi\psi)$ in (7), appears in the second quadrant in Fig. 9b of [1] as $\phi c\phi\tilde{t}^*\sin(-\phi\psi)$.

And a comparison of systems (5) and (3) gives the relations for the induced primed intrinsic affine spacetime coordinates $\phi c\phi\tilde{t}'_i$ and $\phi\tilde{x}'_i$ as follows

$$\phi c\phi\tilde{t}'_i = \phi\tilde{x}'\sin\phi\psi = \frac{\phi v}{\phi c}\phi\tilde{x}'; \quad (8)$$

w.r.t. $3-$ observer in $\tilde{\Sigma}'$ and

$$\phi\tilde{x}'_i = \phi c\phi\tilde{t}'\sin\phi\psi = \frac{\phi v}{\phi c}\phi c\phi\tilde{t}' = \phi v\phi\tilde{t}'; \quad (9)$$

w.r.t. $1-$ observer in $c\tilde{t}'$.

Diagrammatically, the induced primed intrinsic affine time coordinate, $\phi c\phi\tilde{t}'_i = \phi\tilde{x}'\sin\phi\psi$ in Eq. (8), appears in the fourth quadrant in Fig. 8b of [1], where it is written as $\phi\tilde{x}'^*\sin\phi\psi$ and the induced primed intrinsic affine space coordinate, $\phi\tilde{x}'_i = \phi c\phi\tilde{t}'\sin\phi\psi$ in (9), appears in the second quadrant in Fig. 8a of [1], where it is written as $\phi c\phi\tilde{t}'^*\sin\phi\psi$.

The intrinsic affine time induction relation (6) states that an intrinsic affine space coordinate $\phi\tilde{x}$ of the unprimed intrinsic frame, which is inclined at negative intrinsic angle $-\phi\psi$ relative to the intrinsic affine space coordinate $\phi\tilde{x}'$ of the primed intrinsic frame, due to the negative intrinsic speed $-\phi v$ of the unprimed intrinsic frame relative to the primed intrinsic frame, projects (or induces) a negative unprimed intrinsic affine time coordinate, $\phi c\phi\tilde{t}_i = \phi\tilde{x}\sin(-\phi\psi) = -\phi\tilde{x}\sin\phi\psi$, along the vertical relative to the 1-observer in $c\tilde{t}$.

The intrinsic affine space induction relation (7) states that an intrinsic affine time coordinate $\phi c\phi\tilde{t}$ of the observer's (or unprimed) intrinsic frame, which is inclined at negative intrinsic angle $-\phi\psi$ relative to the intrinsic affine time coordinate $\phi c\phi\tilde{t}'$ of the particle's (or primed) intrinsic frame along the vertical, due to the negative intrinsic speed $-\phi v$ of the intrinsic observer's frame relative to the intrinsic particle's frame, induces a negative unprimed intrinsic affine space coordinate, $\phi\tilde{x}_i = \phi c\phi\tilde{t}\sin(-\phi\psi) = -\phi c\phi\tilde{t}\sin\phi\psi$, along the horizontal relative to 3-observer in $\tilde{\Sigma}$.

The intrinsic affine time induction relation (8) states that an intrinsic affine space coordinate $\phi\tilde{x}'$ of the particle's (or primed) intrinsic frame, which is inclined relative to the intrinsic affine space coordinate $\phi\tilde{x}$ of the observer's (or unprimed) intrinsic frame at a positive intrinsic angle $\phi\psi$, due to the intrinsic motion of the particle's (or primed) intrinsic frame at positive intrinsic speed ϕv relative to the observer's (or unprimed) intrinsic frame, induces positive primed intrinsic affine time coordinate, $\phi c\phi\tilde{t}'_i = \phi\tilde{x}'\sin\phi\psi$, along the vertical relative to the 3-observer in $\tilde{\Sigma}'$.

Finally the intrinsic affine space induction relation (9) states that an intrinsic affine time coordinate $\phi c\phi\tilde{t}'$ of the primed intrinsic frame, which is inclined at positive intrinsic angle $\phi\psi$ relative to the intrinsic affine time coordinate $\phi c\phi\tilde{t}$ along the vertical of the primed intrinsic frame, due to the intrinsic motion of the primed intrinsic frame at positive intrinsic speed ϕv relative to the unprimed intrinsic frame, induces positive primed intrinsic affine space coordinate, $\phi\tilde{x}'_i = \phi c\phi\tilde{t}'\sin\phi\psi$, along the horizontal relative to the 1-observer in $c\tilde{t}'$.

The outward manifestations on flat four-dimensional affine spacetime of the intrinsic affine spacetime induction relations (6)–(9) are given by simply removing the symbol ϕ from those relations respectively as follows

$$c\tilde{t}_i = \tilde{x}\sin(-\psi) = -\tilde{x}\sin\psi = -\frac{v}{c}\tilde{x}; \quad (10)$$

w.r.t. $1-$ observer in $c\tilde{t}$;

$$\tilde{x}_i = c\tilde{t}\sin(-\psi) = -c\tilde{t}\sin\psi = -\frac{v}{c}c\tilde{t} = -v\tilde{t}; \quad (11)$$

w.r.t. $3-$ observer in $\tilde{\Sigma}$;

$$c\tilde{t}'_i = \tilde{x}'\sin\psi = \frac{v}{c}\tilde{x}'; \quad (12)$$

w.r.t. $3-$ observer in $\tilde{\Sigma}'$ and

$$\tilde{x}'_i = c\tilde{t}'\sin\psi = \frac{v}{c}c\tilde{t}' = v\tilde{t}'; \quad (13)$$

w.r.t. $1-$ observer in $c\tilde{t}'$.

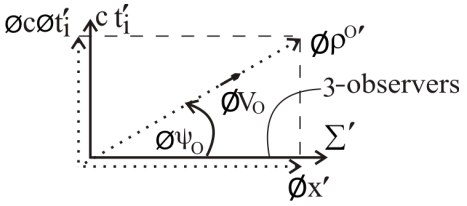


Fig. 7: Proper intrinsic metric time dimension and proper metric time dimension are induced along the vertical with respect to 3-observers in the proper Euclidean 3-space Σ' (as a hyper-surface represented by a line) along the horizontal, by a proper intrinsic metric space that is inclined to the horizontal.

However the derivation of the intrinsic affine spacetime induction relations (6)–(9) in the context of ϕSR and their outward manifestations namely, the affine spacetime induction relations (10)–(13) in the context of SR, are merely to demonstrate explicitly the concept of intrinsic affine spacetime induction that is implicit in intrinsic Lorentz transformation (ϕLT) and its inverse in the context of ϕSR and affine spacetime induction that is implicit in Lorentz transformation (LT) and its inverse in the context of SR.

On the other hand, our interest in this sub-section is in intrinsic metric time induction that arises by virtue of possession of absolute intrinsic speed ϕV_0 naturally at every point along the length of a proper intrinsic metric space, $\phi\rho^{0\prime}$, say, relative to another proper intrinsic metric space, $\phi\rho'$, say, in Fig. 3. Let us assume, for the purpose of illustration, that a proper intrinsic metric space $\phi\rho^{0\prime}$ possesses an absolute intrinsic speed $\phi V_0 < \phi c$ naturally at every point along its length relative to our proper intrinsic metric space $\phi\rho'$ along the horizontal, instead of the absolute intrinsic speed $\phi V_0 = \phi c$ of every point along the length of $\phi\rho^{0\prime}$ relative to $\phi\rho'$ in Fig. 3. Then $\phi\rho^{0\prime}$ will be inclined at an absolute intrinsic angle $\phi\psi_0 < \frac{\pi}{2}$ relative to $\phi\rho'$ along the horizontal, as illustrated in Fig. 7, instead of inclination of $\phi\rho^{0\prime}$ to the horizontal by absolute intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$ in Fig. 3.

As shown in Fig. 7, the inclined proper intrinsic metric space $\phi\rho^{0\prime}$ induces proper intrinsic metric time dimension $\phi c\phi t'_i$ along the vertical with respect to 3-observers in our proper Euclidean 3-space Σ' along the horizontal. The intrinsic metric time induction relation with respect to 3-observers in Σ' in Fig. 7, takes the form of the primed intrinsic affine time induction relation (8) with respect to 3-observer in $\tilde{\Sigma}'$ in the context of ϕSR. We must simply replace the primed intrinsic affine spacetime coordinates $\phi c\phi\tilde{t}'_i$ and $\phi\tilde{x}'$ by proper intrinsic metric spacetime dimensions $\phi c\phi t'_i$ and $\phi\rho^{0\prime}$ respectively and the relative intrinsic angle $\phi\psi$ and relative intrinsic speed ϕv by absolute intrinsic angle $\phi\psi_0$ and absolute intrinsic speed ϕV_0 in (8) to have as follows

$$\phi c\phi t'_i = \phi\rho^{0\prime}\sin\phi\psi_0 = \frac{\phi V_0}{\phi c}\phi\rho^{0\prime}; \qquad (14)$$

w.r.t. all 3 − observers in Σ'. And the outward manifestation of (14) is

$$ct'_i = \rho^{0\prime}\sin\psi_0 = \frac{V_0}{c}\rho^{0\prime}; \qquad (15)$$

w.r.t. all 3 − observers in Σ'.

The induced proper intrinsic metric time dimension $\phi c\phi t'_i$ along the vertical in (14) is made manifest in induced proper metric time dimension ct'_i in (15) along the vertical, as shown in Fig. 7. As indicated, relations (14) and (15) are valid with respect to all 3-observers in our proper Euclidean 3-space Σ' overlying our proper intrinsic metric space $\phi\rho'$ along the horizontal in Fig. 7.

As abundantly stated in [1] and under systems (2) and (3) earlier in this paper, the relative intrinsic angle $\phi\psi = \frac{\pi}{2}$ corresponding to relative intrinsic speed $\phi v = \phi c$, is prohibited by the intrinsic Lorentz transformation (2) and its inverse (3) in ϕSR and consequently $\phi\psi = \frac{\pi}{2}$ or $\phi v = \phi c$ is prohibited in the intrinsic affine time and intrinsic affine space induction relations (6) and (7) and their inverses (8) and (9) in ϕSR. Correspondingly, the angle $\psi = \frac{\pi}{2}$ or speed $v = c$ is prohibited in the affine time and affine space induction relations (10) and (11) and their inverses (12) and (13) in SR.

On the other hand, the absolute intrinsic speed ϕV_0 can be set equal to ϕc and hence the absolute intrinsic angle $\phi\psi_0$ can be set equal to $\frac{\pi}{2}$ in (14). This is so since, as prescribed earlier in this paper, the proper intrinsic metric space $\phi\rho^{0\prime}$ exists naturally along the vertical as in Fig. 3, corresponding to $\phi V_0 = \phi c$ and $\phi\psi_0 = \frac{\pi}{2}$ naturally in (14) with respect to 3-observers in Σ'. More over, as mentioned at the end of subsection 1.1 and as shall be developed fully elsewhere, the absolute intrinsic speed ϕV_0 of every point of the inclined $\phi\rho^{0\prime}$ with respect to all 3-observers in Σ' in Fig. 7, being the outward manifestation in $\phi\rho^{0\prime}$ of the absolute speed of the Newtonian absolute space (the ether of classical mechanics), it can take on values in the range $0 \leq \phi V_0 \leq \infty$, since the maximum speed of objects in classical mechanics is infinite speed. Thus by letting $\phi V_0 = \phi c$ and $\phi\psi_0 = \frac{\pi}{2}$ in (14) we have

$$\phi c\phi t'_i \equiv \phi c\phi t' = \phi\rho^{0\prime}; \quad \text{for } \phi V_0 = \phi c \text{ or } \phi\psi_0 = \frac{\pi}{2} \text{ in Fig. 7;} \qquad (16)$$

w.r.t. all 3 − observers in Σ'.

While relation (14) states that a proper intrinsic metric space $\phi\rho^{0\prime}$, which is inclined to $\phi\rho'$ along the horizontal at absolute intrinsic angle $\phi\psi_0 < \frac{\pi}{2}$, induces proper intrinsic metric time dimension $\phi c\phi t'_i$ along the vertical, whose length is a fraction $\phi V_0/\phi c$ or $\sin\phi\psi_0$ times the length of $\phi\rho^{0\prime}$, with respect to all 3-observers in our proper Euclidean 3-space Σ' along the horizontal, relation (16) states that a proper intrinsic metric space $\phi\rho^{0\prime}$, which is naturally inclined at intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$ relative to $\phi\rho'$ along the horizontal, thereby lying along the vertical, induces proper intrinsic metric time dimension $\phi c\phi t'_i \equiv \phi c\phi t'$ along the vertical, whose length is the

length of $\phi\rho^{0\prime}$, with respect to all 3-observers in our proper Euclidean 3-space Σ' along the horizontal.

The preceding paragraph implies that a proper intrinsic metric space $\phi\rho^{0\prime}$ that is naturally rotated along the vertical is wholly converted (or wholly transformed) into proper intrinsic metric time dimension $\phi c\phi t'$ relative to all observers in our Euclidean 3-space Σ' (along the horizontal). Eq. (16) can therefore be re-written as the transformation of proper intrinsic metric space into proper intrinsic metric time dimension:

$$\phi\rho^{0\prime} \rightarrow \phi c\phi t'; \quad \text{for } \phi V_0 = \phi c \text{ or } \phi\psi_0 = \pi/2 \text{ in Fig. 7;} \quad (17)$$

w.r.t. all 3 − observers in Σ' and the outward manifestation of Eq. (17) is the transformation of the one-dimensional proper metric space $\rho^{0\prime}$ into proper metric time dimension ct':

$$\rho^{0\prime} \rightarrow ct'; \quad \text{for } V_0 = c \text{ or } \psi_0 = \pi/2 \text{ in Eq.(15);} \quad (18)$$

w.r.t. all 3 − observers in Σ'.

The condition required for the transformations (17) and (18) to obtain are naturally met by $\phi\rho^{0\prime}$ and $\rho^{0\prime}$ in Fig. 6a. This is the fact that they are naturally inclined at absolute intrinsic angle $\phi\psi_0 = \frac{\pi}{2}$ and absolute angle $\psi_0 = \frac{\pi}{2}$ respectively relative to our proper intrinsic space $\phi\rho'$ along the horizontal and consequently they naturally possess absolute intrinsic speed $\phi V_0 = \phi c$ and absolute speed $V_0 = c$ respectively at every point along their lengths with respect to all 3-observers in our proper Euclidean 3-space Σ' in that diagram.

The transformations (17) and (18) with respect to 3-observers in our proper Euclidean 3-space Σ' correspond to the following with respect to 3-observers in the proper Euclidean 3-space $-\Sigma'^{*}$ of the negative universe in Fig. 6a:

$$-\phi\rho^{0\prime *} \rightarrow -\phi c\phi t'^{*}; \quad \text{for } \phi V_0 = \phi c \text{ or } \phi\psi_0 = \pi/2; \quad (19)$$

w.r.t. all 3 − observers in $-\Sigma'^{*}$ and

$$-\rho^{0\prime *} \rightarrow -ct'^{*}; \quad \text{for } V_0 = c \text{ or } \psi_0 = \pi/2; \quad (20)$$

w.r.t. all 3 − observers in $-\Sigma'^{*}$.

The counterparts of transformations (17) and (18), which are valid with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ in Fig. 6b are the following

$$\phi\rho' \rightarrow \phi c\phi t^{0\prime}; \quad \text{for } \phi V_0 = \phi c \text{ or } \phi\psi_0 = \pi/2; \quad (21)$$

w.r.t. all 3 − observers in $\Sigma^{0\prime}$ and

$$\rho' \rightarrow ct^{0\prime}; \quad \text{for } V_0 = c \text{ or } \psi_0 = \pi/2; \quad (22)$$

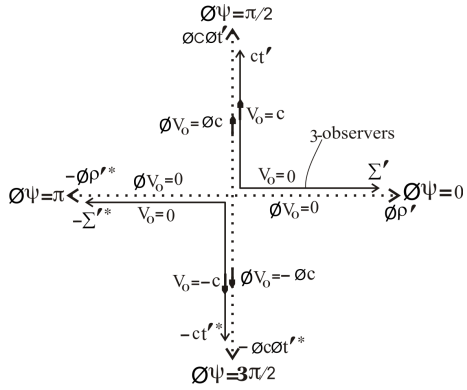


Fig. 8: a) The one-dimensional proper spaces $\rho^{0\prime}$ and $-\rho^{0\prime *}$ in Fig. 6a transform into proper time dimensions ct' and $-ct'^{*}$ respectively and the proper intrinsic spaces $\phi\rho^{0\prime}$ and $-\phi\rho^{0\prime *}$ in Fig. 6a transform into proper intrinsic time dimensions $\phi c\phi t'$ and $-\phi c\phi t'^{*}$ respectively, relative to 3-observers in the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ (represented by lines) along the horizontal.

w.r.t. all 3 − observers in $\Sigma^{0\prime}$ and the counterparts of transformations (19) and (20), which are valid with respect to 3-observers in the proper Euclidean 3-space $-\Sigma^{0\prime *}$ in Fig. 6b are the following

$$-\phi\rho'^{*} \rightarrow -\phi c\phi t^{0\prime *}; \quad \text{for } V_0 = c \text{ or } \psi_0 = \pi/2; \quad (23)$$

w.r.t. all 3 − observers in $-\Sigma^{0\prime *}$ and

$$-\rho'^{*} \rightarrow -ct^{0\prime *}; \quad \text{for } V_0 = c \text{ or } \psi_0 = \pi/2; \quad (24)$$

w.r.t. all 3 − observers in $-\Sigma^{0\prime *}$.

Application of transformations (17)–(20) on Fig. 6a gives Fig. 8a and application of transformation (21)–(24) on Fig. 6b gives Fig. 8b. Again representation of Euclidean 3-spaces by plane surfaces in the previous diagrams in this paper has temporarily been changed to lines in Figs. 8a and 8b, as done in Figs. 6a and 6b, for convenience.

The three-dimensional rest masses of the symmetry-partner particles or objects in the proper Euclidean 3-spaces and the one-dimensional rest masses in the proper time dimensions, as well as their underlying one-dimensional intrinsic rest masses in the proper intrinsic spaces and proper intrinsic time dimensions have been deliberately left out in Figs. 8a and 8b, unlike in Figs. 6a and 6b where they are shown. This is necessary because of further discussion required in locating the one-dimensional particles or objects in the time dimensions, which shall be done later in this paper. As indicated in Figs. 8a and 8b, the proper time dimensions ct' and $ct^{0\prime}$ possess absolute speed $V_0 = c$ at every point along their lengths, relative to 3-observers in the proper Euclidean 3-spaces Σ' and $\Sigma^{0\prime}$ respectively, like the one-dimensional

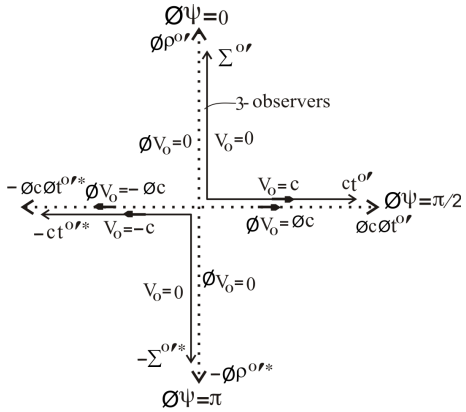


Fig. 8: b) The one-dimensional proper spaces ρ' and $-\rho'^*$ in Fig. 6b transform into proper time dimensions $ct^{0\prime}$ and $-ct^{0\prime *}$ respectively and the proper intrinsic spaces $\phi\rho'$ and $-\phi\rho'^*$ in Fig. 6b transform into proper intrinsic time dimensions $\phi c\phi t^{0\prime}$ and $-\phi c\phi t^{0\prime *}$ respectively relative to 3-observers in the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ (represented by lines) along the vertical.

spaces $\rho^{0\prime}$ and ρ' in Figs. 6a and 6b that transform into ct' and $ct^{0\prime}$ respectively in Figs. 8a and 8b. As also indicated in Figs. 8a and 8b, the proper intrinsic time dimensions $\phi c\phi t'$ and $\phi c\phi t^{0\prime}$ possess absolute intrinsic speed $\phi V_0 = \phi c$ at every point along their lengths relative to 3-observers in the proper Euclidean 3-spaces Σ' and $\Sigma^{0\prime}$ respectively, like the intrinsic spaces $\phi\rho^{0\prime}$ and $\phi\rho'$ in Figs. 6a and 6b that transform into $\phi c\phi t'$ and $\phi c\phi t^{0\prime}$ respectively in Figs. 8a and 8b. The time dimensions ct' and $ct^{0\prime}$ and the intrinsic time dimensions $\phi c\phi t'$ and $\phi c\phi t^{0\prime}$ are isotropic dimensions (with no unique orientations in the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and Σ' that transform into ct' and $ct^{0\prime}$ respectively). These follow from the isotropy of the one-dimensional spaces $\rho^{0\prime}$ and ρ' in the Euclidean 3-spaces $\Sigma^{0\prime}$ and Σ' respectively in Figs. 6a and 6b that transform into ct' and $ct^{0\prime}$ respectively in Figs. 8a and 8b and from the isotropy of the one-dimensional intrinsic spaces $\phi\rho^{0\prime}$ and $\phi\rho'$ in the Euclidean 3-spaces $\Sigma^{0\prime}$ and Σ' respectively in Figs. 6a and 6b that transform into $\phi c\phi t'$ and $\phi c\phi t^{0\prime}$ respectively in Figs. 8a and 8b.

Fig. 8a is the final form to which the quartet of mutually orthogonal proper Euclidean 3-spaces and underlying one-dimensional proper intrinsic spaces in Fig. 2 naturally simplify with respect to 3-observers in the proper Euclidean 3-spaces Σ' and $-\Sigma'^*$ of the positive (or our) universe and the negative universe and Fig. 8b is the final form to which the quartet of proper Euclidean 3-spaces and underlying one-dimensional proper intrinsic spaces in Fig. 2 naturally simplify with respect to 3-observers in the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ of the positive and negative time-universes.

It follows from the natural simplification of Fig. 2 to Figs. 8a and 8b that the concept of time is secondary to the concept of space. Indeed the concept of time had evolved from the concept of space and the concept of intrinsic time had evolved from the concept of intrinsic space. This is so since given the quartet of mutually orthogonal proper metric Euclidean 3-spaces/underlying one-dimensional proper intrinsic metric spaces in Fig. 2, then the straight line one-dimensional proper metric time manifolds (or proper metric time dimensions) evolve automatically relative to 3-observers in the proper Euclidean 3-spaces, as illustrated in Figs. 8a and 8b. Thus one could ask for the origin of space without at the same time asking for the origin of time in the present picture. The origin of time and intrinsic time dimensions, which we seek in this section, has been achieved.

2 Perfect symmetry of natural laws among the isolated four universes

The four universes encompassed by Figs. 8a and 8b are the positive (or our) universe with flat proper spacetime (Σ', ct') of SR and its underlying flat two-dimensional proper intrinsic spacetime $(\phi\rho', \phi c\phi t')$ of ϕSR in Fig. 8a and the negative universe with flat proper spacetime $(-\Sigma'^*, -ct'^*)$ of SR and its underlying two-dimensional flat proper intrinsic spacetime $(-\phi\rho'^*, -\phi c\phi t'^*)$ of ϕSR in Fig. 8a.

The third universe is the one with flat proper spacetime $(\Sigma^{0\prime}, ct^{0\prime})$ of SR and its underlying flat proper intrinsic spacetime $(\phi\rho^{0\prime}, \phi c\phi t^{0\prime})$ of ϕSR in Fig. 8b. This third universe shall be referred to as the positive time-universe, since its proper Euclidean 3-space $\Sigma^{0\prime}$ and its proper intrinsic space $\phi\rho^{0\prime}$ are the proper time dimension ct' and proper intrinsic time dimension $\phi c\phi t'$ respectively of the positive (or our) universe.

The fourth universe is the one with flat proper spacetime $(-\Sigma^{0\prime *}, -ct^{0\prime *})$ of SR and its underlying flat proper intrinsic spacetime $(-\phi\rho^{0\prime *}, -\phi c\phi t^{0\prime *})$ of ϕSR in Fig. 8b. This fourth universe shall be referred to as the negative time-universe, since its proper Euclidean 3-space $-\Sigma^{0\prime *}$ and its proper intrinsic space $-\phi\rho^{0\prime *}$ are the proper time dimension $-ct'^*$ and proper intrinsic time dimension $-\phi c\phi t'^*$ respectively of the negative universe.

As prescribed earlier in this paper, the four worlds (or universes) encompassed by Figs. 8a and 8b, listed above, coexist in nature and exhibit perfect symmetry of natural laws and perfect symmetry of state among themselves. Perfect symmetry of laws among the four universes shall be demonstrated hereunder, while perfect symmetry of state among the universes shall be be demonstrated in the second part of this paper.

Demonstration of perfect symmetry of natural laws between the positive (or our) universe and the negative universe in [1] and [2] involves three steps. In the first step, the affine spacetime/intrinsic affine spacetime diagrams of Figs. 8a and 8b and Figs. 9a and 9b of [1] are derived upon the metric spacetimes/intrinsic metric spacetimes of the positive (or our) universe and the negative universe of Fig. 8a above, (which was prescribed to exist in nature and constitute a two-world background of SR in [1]).

Physical quantity or constant	Symbol	Intrinsic quantity or constant	Sign	
			positive time-universe	negative time-universe
Distance (or dimension) of space	dx^0 or x^0	$d\phi x^0$ or ϕx^0	+	−
Interval (or dimension) of time	dt^0 or t^0	$d\phi t^0$ or ϕt^0	+	−
Mass	m^0	ϕm^0	+	−
Electric charge	q	q	+ or −	− or +
Absolute entropy	S^0	ϕS^0	+	−
Absolute temperature	T	T	+	+
Energy (total, kinetic)	E^0	ϕE^0	+	−
Potential energy	U^0	ϕU^0	+ or −	− or +
Radiation energy	$h\nu^0$	$h\phi\nu^0$	+	−
Electrostatic potential	Φ_E^0	$\phi\Phi_E^0$	+ or −	+ or −
Gravitational potential	Φ^0	$\phi\Phi^0$	−	−
Electric field	$\vec{E}^0$	ϕE^0	+ or −	− or +
Magnetic field	$\vec{B}^0$	ϕB^0	+ or −	− or +
Planck constant	h	h	+	+
Boltzmann constant	k	ϕk	+	−
Thermal conductivity	k	ϕk	+	−
Specific heat capacity	c_p	ϕc_p	+	+
velocity	$\vec{v}$	ϕv	+ or −	+ or −
speeds of particles	v	ϕv	+	+
Speed of light	c	ϕc	+	+
Electric permittivity	ϵ_o^o	$\phi\epsilon_o^o$	+	+
Magnetic permeability	μ_o^o	$\phi\mu_o^o$	+	+
Angular measure	θ, φ	$\phi\theta, \phi\varphi$	+ or −	+ or −
Parity	Π	$\phi\Pi$	+ or −	− or +
⋮	⋮	⋮	⋮	⋮

Table 1. Signs of spacetime/intrinsic spacetime dimensions, some physical parameters/intrinsic parameters and some physical constants/intrinsic constants in the positive time-universe and negative time-universe.

The intrinsic Lorentz transformation/Lorentz transformation (ϕLT/LT) was then derived from those diagrams in the positive and negative universes, thereby establishing intrinsic Lorentz invariance (ϕLI) on flat two-dimensional intrinsic spacetimes and Lorentz invariance (LI) on flat four-dimensional spacetimes in the two universes in [1].

The first step in demonstrating perfect symmetry of laws between the positive (or our) universe and the negative universe in Fig. 8a of this paper described above, applies directly between the positive time-universe and the negative time-universe. The counterparts of Figs. 8a, 8b, 9a and 9b of [1], drawn upon the metric spacetimes/intrinsic metric spacetimes of the positive and negative universes of Fig. 8a of this paper in that paper, can be drawn upon the metric spacetimes/intrinsic spacetimes of the positive time-universe and negative time-universe in Figs. 8b of this paper and intrinsic Lorentz transformations/Lorentz transformation (ϕLT/LT) derived from them in the positive time-universe and the negative time-universe, as shall not be done here in order to conserve space. Intrinsic Lorentz invariance (ϕLI) on flat two-dimensional intrinsic spacetimes and Lorentz invariance (LI) on flat four-dimensional spacetimes in the positive and negative time-universes then follow with respect to observers in those universes.

The second step in demonstrating the symmetry of laws between the positive (or our) universe and the negative universe in [1] and [2], involves the derivation of the relative signs of physical parameters and physical constants and of intrinsic parameters and intrinsic constants between the positive and negative universes in [2], summarized in Table 1 of that paper. Again this second step applies directly between the positive time-universe and the negative time-universe. The relative signs of physical parameters and physical constants and of intrinsic parameters and intrinsic constants derivable between the positive time-universe and the negative time-universe, summarized in Table 1 here, follow directly from the derived signs of physical parameters and physical constants and of intrinsic parameters and intrinsic constants in the pos-

itive and negative universes, summarized in Table 1 of [2].

Table 1 here is the same as Table 1 in [2]. The superscript "0" that appears on dimensions/intrinsic dimensions and some parameters/intrinsic parameters and constants/intrinsic constants in Table 1 here is used to differentiate the dimensions/intrinsic dimensions, parameters/intrinsic parameters and constants/intrinsic constants of the positive time-universe and negative time-universe from those of the positive (or our) universe and the negative universe in Table 1 of [2].

The third and final step in demonstrating the symmetry of natural laws between the positive (or our) universe and the negative universe in [1] and [2], consists in replacing the positive spacetime dimensions and the physical parameters and physical constants that appear in (the instantaneous differential) natural laws in the positive universe by the negative spacetime dimensions and physical parameters and physical constants of the negative universe (with the appropriate signs in Table 1 of [2]), and showing that these operations leave all natural laws unchanged in the negative universe, as done in section 5 of [2].

The third step in the demonstration of the perfect symmetry of natural laws between the positive and negative universes described in the foregoing paragraph, applies directly between the positive time-universe and the negative time-universe as well. Having established Lorentz invariance between the positive time-universe and negative time-universe at the first step, it is straight forward to use Table 1 above and follow the procedure in section 5 of [2] to demonstrate the invariance of natural laws between the positive time-universe and negative time-universe.

Symmetry of natural laws must be considered to have been established between the positive time-universe and the negative time-universe. A more detailed presentation than done above will amount to a repetition of the demonstration of symmetry of natural laws between the positive and negative universes in [1] and [2].

Finally the established validity of Lorentz invariance in the four universes encompassed by Figs. 8a and 8b, coupled with the identical signs of spacetime dimensions, physical parameters and physical constants in the positive (or our) universe and the positive time-universe and the identical signs of spacetime dimensions, physical parameters, physical constants in the negative universe and negative time-universe in Table 1 of [2] and Table 1 above, guarantee the invariance of natural laws between the positive (or our) universe and the positive time-universe and between the negative universe and the negative time-universe. This along with the established invariance of natural laws between the positive (or our) universe and the negative universe and between the positive time-universe and the negative time-universe, guarantees invariance of natural laws among the four universes.

Symmetry of natural laws among the four universes encompassed by Figs. 8a and 8b of this paper namely, the positive (or our) universe and the negative universe (in Fig. 8a), the positive time-universe and the negative time-universe (in Fig. 8b), has thus been shown. Perfect symmetry of state among the universes shall be demonstrated in the second part of this paper, as mentioned earlier.

3 Origin of one-dimensional particles, objects and observers in the time dimension and (3+1)-dimensionality of particles, objects and observers in special relativity

An implication of the geometrical contraction of the three dimensions $x^{01\prime}$, $x^{02\prime}$ and $x^{03\prime}$ of the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe in Fig. 2 or Fig. 3 into a one-dimensional space $\rho^{0\prime}$ relative to 3-observers in our proper Euclidean 3-space Σ' in Fig. 6a, which ultimately transforms into the proper time dimension ct' relative to 3-observers in Σ' in Fig. 8a, is that the dimensions of a particle or object, such as a box of rest mass m_0^0 and proper (or classical) dimensions $\Delta x^{0\prime}$, $\Delta y^{0\prime}$ and $\Delta z^{0\prime}$ in $\Sigma^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$, are geometrically "bundled" parallel to one another, thereby effectively becoming a one-dimensional box of equal rest mass m_0^0 and proper (or classical) length $\Delta\rho^{0\prime}$ along $\rho^{0\prime}$ in Fig. 6a, which transforms into an interval $c\Delta t'$ containing rest mass m_0^0 along the proper time dimension ct' in Fig. 8a, relative to 3-observers in our Euclidean 3-space Σ', where $c\Delta t' = \Delta\rho^{0\prime} = \sqrt{(\Delta x^{0\prime})^2 + (\Delta y^{0\prime})^2 + (\Delta z^{0\prime})^2}$.

Likewise all radial directions of a spherical particle or object of rest mass m_0^0 and proper (or classical) radius $r^{0\prime}$ in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe, with respect to 3-observers in $\Sigma^{0\prime}$, are "bundled" parallel to one another, thereby becoming a one-dimensional particle or object of proper (or classical) length, $\Delta\rho^{0\prime} = r^{0\prime}$, along $\rho^{0\prime}$ in Fig. 6a, which ultimately transforms into interval $c\Delta t'$ $(= r^{0\prime})$ containing rest mass m_0^0 along the proper time dimension ct' in Fig. 8a, with respect to 3-observers in our proper Euclidean 3-space Σ'.

A particle or object of rest mass m_0^0 with arbitrary shape located in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$, will have the lengths (or dimensions) from its centroid to its boundary along all directions geometrically "bundled" parallel to one another, thereby effectively becoming a one-dimensional particle or object of equal rest mass m_0^0 along the proper time dimension ct' with respect to 3-observers in Σ' in Fig. 8a.

The one-dimensional rest mass m_0^0 of proper length $c\Delta t'$ of a particle, object or observer in our proper time dimension ct' with respect to 3-observers in our Euclidean 3-space Σ' in Fig. 8a, will acquire the absolute speed $V_0 = c$, which the proper time dimension possesses at every point along its length with respect to 3-observers in Σ'. Consequently it will possess energy $m_0^0 V_0^2 = m_0^0 c^2 = E'$ in ct' with respect to 3-observers in Σ'. Indeed the one-dimensional rest mass m_0^0 in ct' will be made manifest in the state of energy $E' = m_0^0 c^2$ by virtue of its absolute speed c in ct' and not in the state

of rest mass m_0^0. In other words, instead of locating one-dimensional rest mass m_0^0 along the proper time dimension ct' in Fig. 8a, as done along the one-dimensional space $\rho^{0\prime}$ in Fig. 6a, we must locate one-dimensional equivalent rest mass E'/c^2 (= m_0) along ct' with respect to 3-observers in Σ', as the symmetry-partner in ct' to the three-dimensional rest mass m_0 in Σ'.

It follows from the foregoing that as the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe in Fig. 2 or 3 is geometrically contracted to one-dimensional space $\rho^{0\prime}$ with respect to 3-observers in our proper Euclidean 3-space Σ' in Fig. 6a, the three-dimensional rest mass m_0^0 in $\Sigma^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 2 or Fig. 3, contracts to one-dimensional rest mass m_0^0 located in the one-dimensional space $\rho^{0\prime}$ with respect to 3-observers in our proper Euclidean 3-space Σ' in Fig. 6a. And as the one-dimensional proper space $\rho^{0\prime}$ in Fig. 6a ultimately transforms into the proper time dimension ct' with respect to 3-observers in our Euclidean 3-space Σ', the one-dimensional rest mass m_0^0 in $\rho^{0\prime}$ transforms into one-dimensional equivalent rest mass E'/c^2, (i.e. $m_0^0 \rightarrow E'/c^2$), located in the proper time dimension ct' in Fig. 8a with respect to 3-observers in our Euclidean 3-space Σ', (although E'/c^2 has not been shown in ct' in Fig. 8a).

It must be noted however that since the speed $V_0 = c$ acquired by the rest mass m_0^0 in the proper time dimension ct' is an absolute speed, which is not made manifest in actual motion (or translation) of m_0^0 along ct', the energy $m_0^0c^2 = E'$ possessed by m_0^0 in ct' is a non-detectable energy in the proper time dimension. Important to note also is the fact that the equivalent rest mass E'/c^2 of a particle or object in the proper time dimension ct' is not an immaterial equivalent rest mass. Rather it is a quantity of matter that possesses inertia (like the rest mass m_0^0) along the proper time dimension. This is so because the speed c in $m_0^0c^2 = E'$, being an absolute speed, is not made manifest in motion of m_0^0 along ct', as mentioned above. On the other hand, the equivalent mass, $m_{0\gamma} = E'_\gamma/c^2 = h\nu_0/c^2$, of a photon is purely immaterial, since the speed c in $m_{0\gamma}c^2 = h\nu_0$ is the speed of actual translation through space of photons and only a purely immaterial particle can attain speed c of actual translation in space or along the time dimension. While the material equivalent rest mass $E'/c^2(\equiv m_0^0)$ in ct' can appear as rest mass in SR, the immaterial equivalent mass $E'_{0\gamma}/c^2$ ($\equiv m_{0\gamma}$) of photon cannot appear in SR.

Illustrated in Fig. 9a are the three-dimensional rest mass m_0 of a particle or object at a point of distance d' from a point of reference or origin in our proper Euclidean 3-space Σ' and the symmetry-partner one-dimensional equivalent rest mass E'/c^2 at the symmetry-partner point of distance $d^{0\prime}$ along the proper time dimension ct' from the point of reference or origin, where the distances d' and $d^{0\prime}$ are equal. The three-dimensional rest mass m_0 in Σ' is underlied by its one-dimensional intrinsic rest mass ϕm_0 in the one-dimensional proper intrinsic space $\phi\rho'$ and the one-dimensional equivalent rest mass E'/c^2 in ct' is underlied by its one-dimensional equivalent intrinsic rest mass $\phi E'/\phi c^2$ in the proper intrinsic time dimension $\phi c\phi t'$ in Fig. 9a.

Fig. 9a pertains to a situation where the three-dimensional rest mass m_0 of the particle or object is at rest relative to the 3-observer in the proper Euclidean 3-space Σ' and consequently its one-dimensional equivalent rest mass E'/c^2 is at rest in the proper time dimension ct' relative to the 3-observer in Σ'. On the other hand, Fig. 9b pertains to a situation where the three-dimensional rest mass m_0 of the particle or object is in motion at a velocity $\vec{v}$ relative to the 3-observer in Σ', thereby becoming the special-relativistic mass, $m = \gamma m_0$ in Σ', relative to the 3-observer in Σ' and consequently the one-dimensional equivalent rest mass E'/c^2 of the particle or object is in motion at speed $v = |\vec{v}|$ in the proper time dimension ct' relative to the 3-observer in Σ', thereby becoming the special-relativistic equivalent mass $E/c^2 = \gamma E'/c^2$ in ct' relative to the 3-observer in Σ'.

The one-dimensional equivalent rest mass E'/c^2 of proper (or classical) length $c\Delta t' = d^{0\prime}$ located at a point in the proper time dimension ct' with respect to 3-observers in the proper Euclidean 3-space Σ' in Fig. 9a, acquires the absolute speed $V_0 = c$ of ct'. However, since the absolute speed $V_0 = c$ of ct' is not made manifest in the flow of ct' with respect to 3-observers in Σ', it is not made manifest in translation of E'/c^2 along ct' with respect to the 3-observers in Σ'. Moreover the equivalent rest mass E'/c^2 possesses zero speed ($v = 0$) of motion in ct' relative to the 3-observer in Σ', just as the rest mass m_0 possesses zero speed of motion in the Euclidean 3-space Σ' relative to the 3-observer in Σ'. Consequently m_0 and E'/c^2 remain stationary at their symmetry-partner locations in Σ' and ct' respectively relative to the 3-observer in Σ' in Fig. 9a.

Likewise the equivalent intrinsic rest mass $\phi E'/\phi c^2$ of proper intrinsic length $\phi c\Delta\phi t' = \phi d^{0\prime}$ located at a point in the proper intrinsic time dimension $\phi c\phi t'$ with respect to 3-observers in the proper Euclidean 3-space Σ' in Fig. 9a, acquires the absolute intrinsic speed $\phi V_0 = \phi c$ of $\phi c\phi t'$. However, since the absolute intrinsic speed ϕc of $\phi c\phi t'$ is not made manifest in the intrinsic flow of $\phi c\phi t'$ with respect to 3-observers in Σ', it is not made manifest in intrinsic translation of $\phi E'/\phi c^2$ along $\phi c\phi t'$ with respect to the 3-observers in Σ'. Moreover the equivalent intrinsic rest mass $\phi E'/\phi c^2$ possesses zero intrinsic speed ($\phi v = 0$) of intrinsic translation in $\phi c\phi t'$ relative to the 3-observer in Σ', just as the intrinsic rest mass ϕm_0 possesses zero intrinsic speed of intrinsic translation in the proper intrinsic space $\phi\rho'$ underlying Σ' relative to the 3-observer in Σ'. Consequently ϕm_0 and $\phi E'/\phi c^2$ remain stationary at their symmetry-partner locations in $\phi\rho'$ and $\phi c\phi t'$ respectively relative to the 3-observer in Σ' in Fig. 9a.

In a situation where the rest mass m_0 of the particle or object is in motion at a velocity $\vec{v}$ in the proper Euclidean 3-space Σ' and the one-dimensional equivalent rest mass E'/c^2

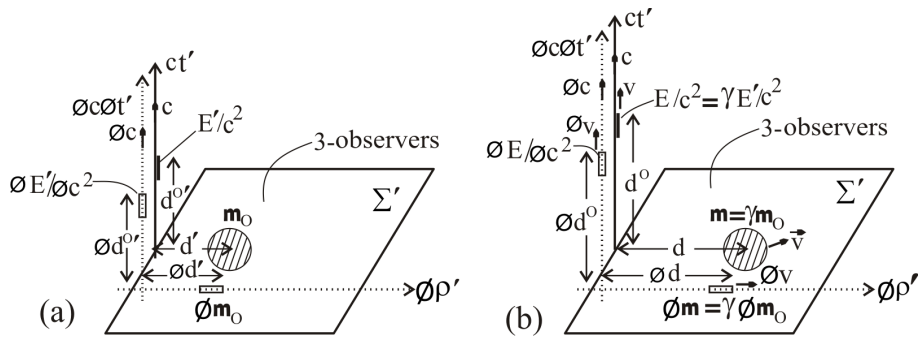


Fig. 9: The three-dimensional mass of an object at a position in the Euclidean 3-space and its one-dimensional equivalent mass at the symmetry-partner position in the time dimension, along with the underlying one-dimensional intrinsic mass of the object in intrinsic space and its equivalent intrinsic mass in the intrinsic time dimension, in the situations where (a) the object is stationary relative to the observer and (b) the object is in motion relative to the observer.

is in motion at speed $v = |\vec{v}|$ in the proper time dimension ct' relative to the 3-observer in Σ' in Fig. 9b, on the other hand, the special-relativistic equivalent mass $E/c^2 = \gamma E'/c^2$, acquires the absolute speed $V_0 = c$ of the proper time dimension ct', which is not made manifest in motion of $\gamma E'/c^2$ along ct' and as well possesses speed v of translation along ct' relative to the 3-observer in Σ'.

During a given period of time, the relativistic equivalent mass $\gamma E'/c^2$ has translated at constant speed v from an initial position P_1^0 to another position P_2^0 along the proper time dimension ct', while covering an interval $P_1^0P_2^0$ of ct'. During the same period of time, the special-relativistic mass $m = \gamma m_0$, has translated at equal constant speed $v = |\vec{v}|$ from an initial position P_1 to another position P_2 in the proper Euclidean 3-space Σ', while covering a distance P_1P_2 in Σ', where the interval $P_1^0P_2^0$ covered along ct' by $\gamma E'/c^2$ is equal to the distance P_1P_2 covered in Σ' by γm_0 and positions P_1 and P_2 in Σ' are symmetry-partner positions to positions P_1^0 and P_2^0 respectively in ct'. Consequently γm_0 and $\gamma E'/c^2$ are always located at symmetry-partner positions in Σ' and ct' respectively in the situation where they are in motion at any speed v in their respective domains relative to the 3-observer in Σ' in Fig. 9b.

It shall be reiterated for emphasis that the equivalent mass E'/c^2 or $\gamma E'/c^2$ in our proper metric time dimension ct' with respect to 3-observers in our proper Euclidean 3-space Σ', of a particle, object or observer in Figs. 9a and 9b, is actually the three-dimensional mass m_0^0 or γm_0^0 of the symmetry-partner particle, object or observer in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$. This is the origin of the the one-dimensional particle, object or observer (or 1-particle, 1-object or 1-observer) in the time dimension to every 3-dimensional particle, object or observers (or 3-particle, 3-object or 3-observer) in 3-space in our universe.

Just as the proper time dimension $ct'(\equiv x^{0\prime})$ is added to the three dimensions $x^{1\prime}$, $x^{2\prime}$ and $x^{3\prime}$ of the proper Euclidean 3-space Σ' to have the four dimensions $x^{0\prime}$, $x^{1\prime}$, $x^{2\prime}$ and $x^{3\prime}$ of the flat four-dimensional proper metric spacetime, the one-dimensional equivalent rest mass E'/c^2 of a particle, object or observer in the proper time dimension ct' must be added to the three-dimensional rest mass m_0 of its symmetry-partner particle, object or observer in the proper Euclidean 3-space Σ' to have a 4-dimensional particle, object or observer of rest mass $(m_0, E'/c^2)$ on the flat four-dimensional proper spacetime (Σ', ct') in our notation.

However it is more appropriate to refer to 4-dimensional particles, objects and observers on flat 4-dimensional spacetime as (3+1)-dimensional particles, objects and observers, because the one-dimensional particles, objects and observers (or 1-particles, 1-objects and 1-observers) in the time dimension ct' are themselves distinct particles, objects and observers, (which are geometrically contracted from three-dimensional particles, objects and observers in the Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe), which are separated in the time dimension ct' from their symmetry-partner three-dimensional continuum particles, objects and observers (or 3-particles, 3-objects and 3-observers) in the continuum Euclidean 3-space Σ'.

The 1-particle, 1-object or 1-observer in the time dimension can be thought of as weakly bonded to the 3-particle, 3-object or 3-observer in the Euclidean 3-space to form a (3+1)-dimensional particle, object or observer in spacetime and a (3+1)-dimensional particle, object or observer can be decomposed into its component 1-particle, 1-object or 1-observer in the time dimension and 3-particle, 3-object or 3-observer in the Euclidean 3-space. On the other hand, what should be referred to as a continuum 4-dimensional particle, object or observer (or 4-particle, 4-object or 4-observer) on four-dimensional spacetime continuum should be non-decomposable into its component dimensions, just as a continuum 3-dimensional particle, object or observer in the Euclidean 3-space continuum cannot be decomposed into its component dimensions.

There are no continuum non-decomposable four-dimensional particles, objects and observers on four-dimensional spacetime in the context of the present theory. Rather there are (3+1)-dimensional particles, objects and observers that can be decomposed into one-dimensional particles, objects and observers in the time dimension and three-dimensional particles, objects and observers in the Euclidean 3-space. Relativistic physics must be formulated partially with respect to 1-observers in the time dimension as distinct from relativistic physics formulated partially with respect to 3-observers in the Euclidean 3-space. The partial physics formulated with respect to 1-observer in the time dimension and 3-observer in the Euclidean 3-space must then be composed into the full relativistic physics on four-dimensional spacetime.

It is also important to note that it is the partial physics formulated with respect to 1-observers in the time dimension, which, of course, contains component of physics projected from the Euclidean 3-space in relativistic physics, is what the 1-observers in the time dimension could observe. It is likewise the partial physics formulated with respect to 3-observers in the Euclidean 3-space, which, of course, contains component of physics projected from the time dimension in relativistic physics, that the 3-observers in the Euclidean 3-space could observe.

The foregoing paragraph has been well illustrated with the derivation of the intrinsic Lorentz transformation of system (13) of [1] as combination of partial intrinsic Lorentz transformation (11) derived from Fig. 8a with respect to the 3-observer (Peter) in the Euclidean 3-space $\tilde{\Sigma}$ and partial intrinsic Lorentz transformation (12) derived from Fig. 8b with respect to the 1-observer ($\tilde{\text{P}}$eter) in the time dimension $c\tilde{t}$ in that paper. The Lorentz transformation of system (28) of [1], as the outward manifestation on flat four-dimensional spacetime of the intrinsic Lorentz transformation (11) in that paper, has likewise been composed from partial Lorentz transformation with respect to the 3-observer in $\tilde{\Sigma}$ and partial Lorentz transformation with respect to the 1-observer in the time dimension $c\tilde{t}$.

Let us collect the partial Lorentz transformations derived with respect to the 1-observer in $c\tilde{t}$ in the LT and its inverse of systems (28) and (29) of [1] to have as follows

$$\left.\begin{array}{rcl} c\tilde{t}' & = & c\tilde{t}\sec\psi - \tilde{x}\tan\psi; \\ \tilde{x} & = & \tilde{x}'\sec\psi + c\tilde{t}\tan\psi;\ \tilde{y} = \tilde{y}';\ \tilde{z} = \tilde{z}'; \\ & & (\text{w.r.t. } 1-\text{observer in } c\tilde{t}) \end{array}\right\}. \qquad (25)$$

These coordinate transformations simplify as follows from the point of view of what can be measured with laboratory rod and clock discussed in detail in sub-section 4.5 of [1]:

$$\tilde{t} = \tilde{t}'\cos\psi;\ \tilde{x} = \tilde{x}'\sec\psi;\ \tilde{y} = \tilde{y}';\ \tilde{z} = \tilde{z}' \qquad (26)$$

w.r.t. $1-$ observer in $c\tilde{t}$.

System (26) derived with respect to the 1-observer in $c\tilde{t}$, corresponds to system (42) of [1], derived with respect to 3-observer in $\tilde{\Sigma}$ in that paper, which shall be re-presented here as follows

$$\tilde{t} = \tilde{t}'\sec\psi;\ \tilde{x} = \tilde{x}'\cos\psi;\ \tilde{y} = \tilde{y}';\ \tilde{z} = \tilde{z}' \qquad (27)$$

w.r.t. $3-$ observer in $\tilde{\Sigma}$.

We find from systems (26) and (27) that while 3-observers in the Euclidean 3-space observe length contraction and time dilation of relativistic events, their symmetry-partner 1-observers in the time dimension observe length dilation and time contraction of relativistic events.

It is clear from all the foregoing that a 3-observer in the Euclidean 3-space and his symmetry-partner 1-observer in the time dimension are distinct observers who can be composed (or "weakly bonded") into a (3+1)-dimensional observer that can be decomposed back into its component 3-observer and 1-observer for the purpose of formulating relativistic physics, which is composed from partial relativistic physics formulated separately with respect to 3-observers in the Euclidean 3-space and 1-observers in the time dimension.

Every parameter in the Euclidean 3-space has its counterpart (or symmetry-partner) in the time dimension. We have seen the case of rest mass m_0 in the proper Euclidean 3-space Σ' and its symmetry-partner one-dimensional equivalent rest mass E'/c^2 in the proper time dimension ct', as illustrated in Figs. 9a and 9b. A classical three-vector quantity $\vec{q}\,'$ in the proper Euclidean 3-space Σ' has its symmetry-partner classical scalar quantity $q^{0\prime}$ in the proper time dimension ct'. The composition of the two yields what is usually referred to as four-vector quantity denoted by $q'_\lambda = (q^{0\prime}, \vec{q}\,')$ or $q'_\lambda = (q^{0\prime}, q^{1\prime}, q^{2\prime}, q^{3\prime})$. We now know that the scalar components $q^{0\prime}$ in the time dimension ct' of four-vector quantities in the positive (or our) universe are themselves three-vector quantities $\vec{q}^{\,0\prime}$ in the Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$. The three-vector quantities $\vec{q}^{\,0\prime}$ in $\Sigma^{0\prime}$, (which are identical symmetry-partners to the three-vector quantities $\vec{q}\,'$ in our Euclidean 3-space Σ'), become contracted to one-dimensional scalar quantities $q^{0\prime} = |\vec{q}^{\,0\prime}|$ in the time dimension ct' relative to 3-observers in Σ', even as the proper Euclidean 3-space $\Sigma^{0\prime}$ containing $\vec{q}^{\,0\prime}$ becomes contracted to the proper time dimension ct' relative to 3-observers in Σ'.

4 Final justification for the new spacetime/intrinsic spacetime diagrams for Lorentz transformation/intrinsic Lorentz transformation in the four-world picture

New geometrical representations of Lorentz transformation and intrinsic Lorentz transformation (LT/ϕLT) and their inverses were derived and presented as Figs. 8a and 8b and Figs. 9a and 9b within the two-world picture isolated in [1]. However at least two outstanding issues about those diagrams remain to be resolved in order to finally justify them. The first issue is the unexplained origin of Fig. 8b that must necessarily be drawn to complement Fig. 8a of [1] in deriving ϕLT/LT.

The second issue is the unspecified reason why anticlockwise relative rotations of intrinsic affine spacetime coordinates are positive rotations (involving positive intrinsic angles $\phi\psi$) with respect to 3-observers in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ in Fig. 8a of [1], while, at the same time, clockwise relative rotations of intrinsic affine spacetime coordinates are positive rotations (involving positive intrinsic angles $\phi\psi$) with respect to 1-observers in the time dimensions ct' and $-ct'^{*}$ in Fig. 8b of [1]. These two issues shall be resolved within the four-world picture encompassed by Figs. 8a and 8b of this paper in this section.

Let us as done in deriving Figs. 8a and 8b and their inverses Figs. 9a and 9b of [1] towards the derivation of intrinsic Lorentz transformation/Lorentz transformation (ϕLT/LT) and their inverses in the positive and negative universes in [1], prescribe particle's (or primed) frame and observer's (or unprimed) frame in terms of extended affine spacetime coordinates in the positive (or our) universe as $(\tilde{x}', \tilde{y}', \tilde{z}', c\tilde{t}')$ and $(\tilde{x}, \tilde{y}, \tilde{z}, c\tilde{t})$ respectively. They are underlied by intrinsic particle's frame and intrinsic observer's frame in terms of extended intrinsic affine coordinates $(\phi\tilde{x}', \phi c\phi\tilde{t}')$ and $(\phi\tilde{x}, \phi c\phi\tilde{t})$ respectively.

The prescribed perfect symmetry of state between the positive and negative universes in [1] implies that there are identical symmetry-partner particle's frame and observer's frame $(-\tilde{x}'^{*}, -\tilde{y}'^{*}, -\tilde{z}'^{*}, -c\tilde{t}'^{*})$ and $(-\tilde{x}^{*}, -\tilde{y}^{*}, -\tilde{z}^{*}, -c\tilde{t}^{*})$ respectively, as well as their underlying identical symmetry-partner intrinsic particle's frame and symmetry-partner intrinsic observer's frame $(-\phi\tilde{x}'^{*}, -\phi c\phi\tilde{t}'^{*})$ and $(-\phi\tilde{x}^{*}, -\phi c\phi\tilde{t}^{*})$ respectively in the negative universe.

Let us consider the motion at a constant speed v of the rest mass m_0 of the particle along the $\tilde{x}'-$axis of its frame and the underlying intrinsic motion at constant intrinsic speed ϕv of the intrinsic rest mass ϕm_0 of the particle along the intrinsic space coordinate $\phi\tilde{x}'$ of its frame relative to a 3-observer in the positive universe. Again the prescribed perfect symmetry of state between the positive and negative universes implies that the rest mass $-m_0^{*}$ of the symmetry-partner particle is in simultaneous motion at equal constant speed v along the $-\tilde{x}'^{*}-$axis of its frame of reference and its intrinsic rest mass $-\phi m_0^{*}$ is in simultaneous intrinsic motion at equal intrinsic speed ϕv along the intrinsic space coordinate $-\phi\tilde{x}'^{*}-$axis of its frame relative to the symmetry-partner 3-observer in the negative universe.

As developed in sub-section 4.4 of [1], the simultaneous identical motions of the symmetry-partner particles' frames relative to the symmetry-partner observers' frames in the positive and negative universes, described in the foregoing paragraph, give rise to Fig. 8a of [1] with respect to 3-observers in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$, which shall be reproduced here as Fig. 10a.

The diagram of Fig. 10a involving relative rotations of extended intrinsic affine spacetime coordinates, has been drawn upon the flat four-dimensional proper metric spacetime of classical mechanics (CM) and its underlying flat two-dimensional proper intrinsic metric spacetime of intrinsic classical mechanics (ϕCM) of the positive (or our) universe and the negative universe contained in Fig. 8a of this paper. The prescribed symmetry of state among the four universes encompassed by Figs. 8a and 8b of this paper, implies that identical symmetry-partner particles undergo identical motions simultaneously relative to identical symmetry-partner observers (or frames of reference) in the four universes. It follows from this that Fig. 10b drawn upon the flat four-dimensional proper metric spacetime of CM and its underlying flat two-dimensional proper intrinsic metric spacetime of ϕCM of the positive time-universe and the negative time-universe contained in Fig. 8b of this paper, co-exists with Fig. 10a in nature.

Fig. 10b is valid with respect to 3-observers in the Euclidean 3-spaces $\Sigma^{0\prime}$ of the positive time-universe and $-\Sigma^{0\prime *}$ of the negative time-universe as indicated. It must be noted that the anti-clockwise rotations of primed intrinsic coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ relative to the unprimed intrinsic coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ respectively by positive intrinsic angle $\phi\psi$ with respect to 3-observers in the Euclidean 3-space Σ' and $-\Sigma'^{*}$ in Fig. 10a, correspond to clockwise rotations of the primed intrinsic coordinates $\phi\tilde{x}^{0\prime}$ and $\phi c\phi\tilde{t}^{0\prime}$ relative to the unprimed intrinsic coordinates $\phi\tilde{x}^{0}$ and $\phi c\phi\tilde{t}^{0}$ respectively by positive intrinsic angle $\phi\psi$ with respect to 3-observers in $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ in Fig. 10b.

Fig. 10b co-exists with Fig. 10a in nature and must complement Fig. 10a towards deriving intrinsic Lorentz transformation/Lorentz transformation (ϕLT/LT) graphically in the positive (or our) universe and the negative universe by physicists in our universe and the negative universe. However Fig. 10b in its present form cannot serve a complementary role to Fig. 10a, because it contains the spacetime and intrinsic spacetime coordinates of the positive time-universe and the negative time-universe, which are elusive to observers in our (or positive) universe and the negative universe, or which cannot appear in physics in the positive and negative universes.

In order for Fig. 10b to be able to serve a complementary role to Fig. 10a towards deriving the ϕLT/LT in the positive and negative universes, it must be appropriately modified. As found earlier in this paper, the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ of the positive and negative time-universes with respect to 3-observers in them, are proper time dimensions ct' and $-ct'^{*}$ respectively with respect to 3-observers in the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ of our universe and the negative universe and the proper time dimensions $ct^{0\prime}$ and $-ct^{0\prime *}$ of the positive and negative time-universes with respect to 3-observers in the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ of the positive and negative time-universes, are the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ of our universe and the negative universes respectively with respect to 3-observers in Σ' and $-\Sigma'^{*}$.

As follows from the foregoing paragraph, Fig. 10b will contain the spacetime and intrinsic spacetime coordinates of

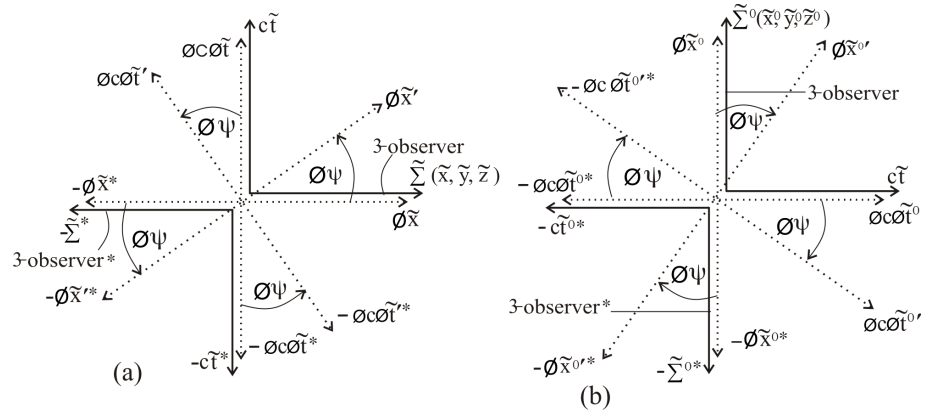


Fig. 10: a) Relative rotations of intrinsic affine spacetime coordinates of a pair of frames in the positive (or our) universe and of the symmetry-partner pair of frames in the negative universe, which are valid relative to symmetry-partner 3-observers in the Euclidean 3-spaces in the positive and negative universes. b) Relative rotations of intrinsic affine spacetime coordinates of a pair of frames in the positive time-universe and of the symmetry-partner pair of frames in the negative time-universe, which are valid relative to symmetry-partner 3-observers in the Euclidean 3-spaces in the positive and negative time-universes.

our (or positive) universe and the negative universe solely by performing the following transformations of spacetime and intrinsic spacetime coordinates on it with respect to 3-observers in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ of our universe and the negative universe:

$$\left.\begin{array}{lcl} \tilde{\Sigma}^0 & \rightarrow & c\tilde{t};\ c\tilde{t}^0 \rightarrow \tilde{\Sigma};\ -\tilde{\Sigma}^{0*} \rightarrow -c\tilde{t}^{*};\\ & & -c\tilde{t}^{0*} \rightarrow -\tilde{\Sigma}^{*}. \\ \phi\tilde{x}^0 & \rightarrow & \phi c\phi\tilde{t};\ \phi c\phi\tilde{t}^0 \rightarrow \phi\tilde{x};\\ & & -\phi\tilde{x}^{0*} \rightarrow -\phi c\phi\tilde{t}^{*};\\ & & -\phi c\phi\tilde{t}^{0*} \rightarrow -\phi\tilde{x}^{*}. \\ \phi\tilde{x}^{0\prime} & \rightarrow & \phi c\phi\tilde{t}';\ \phi c\phi\tilde{t}^{0\prime} \rightarrow \phi\tilde{x}';\\ & & -\phi\tilde{x}^{0\prime *} \rightarrow -\phi c\phi\tilde{t}'^{*};\\ & & -\phi c\phi\tilde{t}^{0\prime *} \rightarrow -\phi\tilde{x}'^{*}. \end{array}\right\} \quad (28)$$

By implementing the coordinate/intrinsic coordinate transformations of systems (28) on Fig. 10b we have Fig. 11a.

Fig. 11a is valid with respect to 1-observers in the proper time dimensions ct' and $-ct'^{*}$ of the positive and negative universes as indicated, where these 1-observers are the 3-observers in the Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ in Fig. 10b. Since Fig. 11a contains the spacetime/intrinsic spacetime coordinates of the positive (or our) universe and the negative universe solely, it can serve as a complementary diagram to Fig. 10a towards the deriving ϕLT/LT in the positive (or our) universe and the negative universe. Indeed Fig. 10a and Fig. 11a are the same as Figs. 8a and 8b of [1], with which the ϕLT/LT were derived in the positive (or our) universe and the negative universe in that paper, except for intrinsic spacetime projections in Figs. 8a and 8b of [1], which are not shown in Figs. 10a and Fig. 11a here.

On the other hand, Fig. 10a will contain the spacetime/intrinsic spacetime coordinates of the positive time-universe and the negative time-universe solely, as shown in Fig. 11b, by performing the inverses of the transformations of spacetime and intrinsic spacetime coordinates of system (28), (that is, by reversing the directions of the arrows in system (28)) on Fig. 10a. Just as Fig. 11a must complement Fig. 10a for the purpose of deriving the ϕLT/LT in the positive (or our) universe and the negative universe, as presented in sub-section 4.4 of [1], Fig. 11b must complement Fig. 10b for the purpose of deriving the ϕLT/LT in the positive time-universe and the negative time-universe.

The clockwise sense of relative rotations of intrinsic affine spacetime coordinates by positive intrinsic angles $\phi\psi$ with respect to 1-observers in the time dimension $c\tilde{t}$ and $-c\tilde{t}^{*}$ in Fig. 11a follows from the validity of the clockwise sense of relative rotations of intrinsic affine spacetime coordinates by positive intrinsic angle $\phi\psi$ with respect to 3-observers in the Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ in Fig. 10b. The 1-observers in $c\tilde{t}$ and $-c\tilde{t}^{*}$ in Fig. 11a are what the 3-observers in $\tilde{\Sigma}^0$ and $-\tilde{\Sigma}^{0*}$ in Fig. 10b transform into, as noted above.

Thus the second outstanding issue about the diagrams of Figs. 8a and 8b of [1], mentioned at the beginning of this section namely, the unexplained reason why anti-clockwise relative rotations of intrinsic affine spacetime coordinates with respect to 3-observers in the Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ are positive rotations involving positive intrinsic angles $\phi\psi$ in Fig. 8a of [1], while, at the same time, clockwise relative rotations of intrinsic affine spacetime coordinates with respect to 1-observers in the time dimensions ct' and $-ct'^{*}$ are positive rotations involving positive intrinsic angles $\phi\psi$ in Fig. 8b of [1], has now been resolved.

Since Fig. 8b of [1] or Fig. 11a of this paper has been shown to originate from Fig. 10b of this paper, which is valid with respect to 3-observers in the Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ of the positive and negative time-universes, the origin

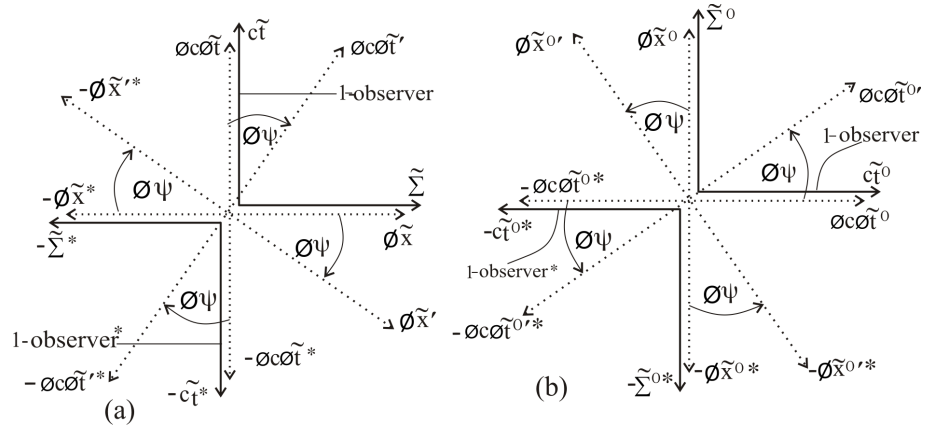


Fig. 11: a) Complementary diagram to Fig. 10a obtained by transforming the spacetime/intrinsic spacetime coordinates of the positive time-universe and the negative time-universe in Fig. 10b into the spacetime/intrinsic spacetime coordinates of the positive (or our) universe and the negative universe; is valid with respect to 1-observers in the time dimensions of our universe and the negative universe. b) Complementary diagram to Fig. 10b obtained by transforming the spacetime/intrinsic spacetime coordinates of the positive (or our) universe and the negative universe in Fig. 10a into the spacetime/intrinsic spacetime coordinates of the positive time-universe and the negative time-universe; is valid with respect to 1-observers in the time dimensions of the positive time-universe and the negative time-universe.

from the positive time-universe and negative time-universe of Fig. 8b of [1] (or Fig. 11a of this paper), which must necessarily be drawn to complement Fig. 8a of [1] (or Fig. 10a of this paper) in deriving the ϕLT/LT in our (or positive) universe and the negative universe, has been shown. Thus the first outstanding issue about Figs. 8a and 8b of [1], which was unresolved in [1], mentioned at the beginning of this section, namely the unexplained origin of Fig. 8b that must always be drawn to complement Fig. 8a in [1] in deriving the ϕLT/LT, has now been resolved. The four-world background of Figs. 8a and its complementary diagram of Fig. 8b in [1] (or Fig. 10a and Fig. 11a of this paper), has thus been demonstrated.

The new geometrical representation of the intrinsic Lorentz transformation/Lorentz transformation (ϕLT/LT) of Figs. 8a and 8b in [1] (or Fig. 10a and Fig. 11a of this paper), which was said to rest on a two-world background in [1] and [2], because those diagrams contain the spacetime/intrinsic spacetime coordinates of the positive (or our) universe and the negative universe solely and the origin of Fig. 8b in [1] (or Fig. 11a of this paper) from the diagram of Fig. 10b of this paper in the positive time-universe and the negative time-universe was unknown in [1]. The ϕLT/LT and consequently the intrinsic special theory of relativity/special theory of relativity (ϕSR/SR) shall be said to rest on a four-world background henceforth.

5 Invariance of the flat four-dimensional proper (or classical) metric spacetime in the context of special relativity

The flat four-dimensional proper physical (or metric) spacetime, which is composed of the proper Euclidean 3-space Σ' and the proper time dimension ct' in the first quadrant in Fig. 8a of this paper, is the flat four-dimensional proper metric spacetime of classical mechanics (including classical gravitation), of the positive (or our) universe, usually denoted by $(x^{0\prime}, x^{1\prime}, x^{2\prime}, x^{3\prime})$, where the dimension $x^{0\prime}$ is along the one-dimensional proper space $\rho^{0\prime}$ in Fig. 6a, which transforms into the proper time dimension ct' in Fig. 8a; hence $x^{0\prime} = ct'$ and $x^{1\prime}$, $x^{2\prime}$ and $x^{3\prime}$ are the dimensions of the proper Euclidean 3-space Σ'. The notation (Σ', ct') for the flat four-dimensional proper physical (or metric) spacetime adopted in [1] and [2], (although the prime label on Σ' and ct' did not appear in those papers), is being adhered to in this paper for convenience.

When the special theory of relativity operates on the flat four-dimensional proper metric spacetime $(x^{0\prime}, x^{1\prime}, x^{2\prime}, x^{3\prime})$; $x^{0\prime} = ct'$ (or (Σ', ct') in our notation), it is the extended intrinsic affine spacetime coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ of the primed (or particle's) frame that are rotated relative to their projective extended affine intrinsic spacetime coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ of the unprimed (or observer's) frame. It is consequently the primed intrinsic affine coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ that transform into the unprimed intrinsic affine coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ in intrinsic Lorentz transformation (ϕLT) in the context of intrinsic special theory of relativity (ϕSR).

It is the extended affine spacetime coordinates $c\tilde{t}'$, $\tilde{x}'$, $\tilde{y}'$ and $\tilde{z}'$ of the primed frame on the flat four-dimensional proper physical (or metric) spacetime $(x^{0\prime}, x^{1\prime}, x^{2\prime}, x^{3\prime})$ (or (Σ', ct') in our notation) that transform into the extended affine spacetime coordinates $c\tilde{t}$, $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ of the unprimed frame, also on the flat four-dimensional proper physical (or metric) spacetime $(x^{0\prime}, x^{1\prime}, x^{2\prime}, x^{3\prime})$ (or (Σ', ct') in our notation) in Lorentz transformation (LT) in the context of the special theory of

relativity (SR).

The special theory of relativity, as an isolated phenomenon, cannot transform the extended flat proper metric spacetime ($x^{0\prime}$, $x^{1\prime}$, $x^{2\prime}$, $x^{3\prime}$) (or (Σ', ct') in our notation) on which it operates, to an extended flat relativistic metric spacetime (x^0, x^1, x^2, x^3) (or (Σ, ct) in our notation), because SR involves the transformation of extended affine spacetime coordinates with no physical (or metric) quality. Or because the spacetime geometry associated with SR is affine spacetime geometry. A re-visit to the discussion of affine and metric spacetimes in sub-section 4.4 of [1] may be useful here. The primed coordinates $\tilde{x}'$, $\tilde{y}'$, $\tilde{z}'$ and $c\tilde{t}'$ of the particle's frame and the unprimed coordinates $\tilde{x}$, $\tilde{y}$, $\tilde{z}$ and $c\tilde{t}$ of the observer's frame in the context of SR are affine coordinates with no metric quality, both of which exist on the flat proper (or classical) metric spacetime ($x^{0\prime}$, $x^{1\prime}$, $x^{2\prime}$, $x^{3\prime}$) (or (Σ', ct') in our notation).

It is gravity (a metric phenomenon) that can transform extended flat four-dimensional proper (or classical) metric spacetime (with prime label) ($x^{0\prime}$, $x^{1\prime}$, $x^{2\prime}$, $x^{3\prime}$) (or (Σ', ct') in our notation) into extended four-dimensional "relativistic" spacetime (x^0, x^1, x^2, x^3) (or (Σ, ct)), (without prime label), where (x^0, x^1, x^2, x^3) (or (Σ, ct)) is known to be curved in all finite neighborhood of a gravitation field source in the context of the general theory of relativity (GR). The rest mass m_0 of a test particle on the flat proper (or classical) metric spacetime ($x^{0\prime}$, $x^{1\prime}$, $x^{2\prime}$, $x^{3\prime}$) (or (Σ', ct')) is also known to transform into the inertial mass m on the curved "relativistic" physical (or metric) spacetime (x^0, x^1, x^2, x^3) (or (Σ, ct)) in the context of GR, where m is known to be trivially related to m_0 as $m = m_0$, by virtue of the principle of equivalence of Albert Einstein [5].

However our interest in [1] and [2] and in the two parts of this paper is not in the metric phenomenon of gravity, but in the special theory of relativity (with affine spacetime geometry), as an isolated subject from gravity. We have inherently assumed the absence of gravity by restricting to the extended flat four-dimensional proper (or classical) metric spacetime ($x^{0\prime}$, $x^{1\prime}$, $x^{2\prime}$, $x^{3\prime}$) (or (Σ', ct') in our notation), as the metric spacetime that supports SR in the absence of relativistic gravity in [1] and [2] and up to this point in this paper. The transformation of the flat proper (or classical) metric spacetime ($x^{0\prime}$, $x^{1\prime}$, $x^{2\prime}$, $x^{3\prime}$) (or (Σ', ct')) into "relativistic" metric spacetime (x^0, x^1, x^2, x^3) (or (Σ, ct)) in the context of a theory of gravity, shall be investigated with further development within the present four-world picture, in which four-dimensional spacetime is underlined by two-dimensional intrinsic spacetime in each of the four symmetrical worlds (or universes).

This first part of this paper shall be ended at this point, while justifications for the co-existence in nature of the four symmetrical worlds (or universes) in Figs. 8a and 8b of this paper, as the actual background of the special theory of relativity in each universe, shall be concluded in the second part.

Submitted on January 26, 2010 / Accepted on March 01, 2010

References

1. Adekugbe A. O. J. Two-world background of special relativity. Part I. *Progress in Physics*, 2010, v. 1, 30–48.
2. Adekugbe A. O. J. Two-world background of special relativity. Part II. *Progress in Physics*, 2010, v. 1, 49–61.
3. Adler R., Bazin M., Schiffer M. Introduction to general relativity. Second Edition. McGraw-Hill Book Company, New York, 1975.
4. Einstein A. Sidelights on relativity. Methuen, 1922.
5. Bonnor W. B. Negative mass in general relativity. *Gen. Relat. Grav.*, 1989, v. 21, 1143–1157.

SPECIAL REPORT

Re-Identification of the Many-World Background of Special Relativity as Four-World Background. Part II.

Akindele O. J. Adekugbe

Center for The Fundamental Theory, P. O. Box 2575, Akure, Ondo State 340001, Nigeria.
E-mail: adekugbe@alum.mit.edu

The re-identification of the many-world background of the special theory of relativity (SR) as four-world background in the first part of this paper (instead of two-wold background isolated in the initial papers), is concluded in this second part. The flat two-dimensional intrinsic spacetime, which underlies the flat four-dimensional spacetime in each universe, introduced as *ansatz* in the initial paper, is derived formally within the four-world picture. The identical magnitudes of masses, identical sizes and identical shapes of the four members of every quartet of symmetry-partner particles or objects in the four universes are shown. The immutability of Lorentz invariance on flat spacetime of SR in each of the four universes is shown to arise as a consequence of the perfect symmetry of relative motion at all times among the four members of every quartet of symmetry-partner particles and objects in the four universes. The perfect symmetry of relative motions at all times, coupled with the identical magnitudes of masses, identical sizes and identical shapes, of the members of every quartet of symmetry-partner particles and objects in the four universes, guarantee perfect symmetry of state among the universes.

1 Isolating the two-dimensional intrinsic spacetime that underlies four-dimensional spacetime

1.1 Indispensability of the flat 2-dimensional intrinsic spacetime underlying flat 4-dimensional spacetime

The flat two-dimensional proper intrinsic metric spacetimes denoted by $(\phi\rho', \phi c\phi t')$ and $(-\phi\rho'^{*}, -\phi c\phi t'^{*})$, which underlies the flat four-dimensional proper metric spacetimes (Σ', ct') and $(-\Sigma'^{*}, -ct'^{*})$ of the positive and negative universes respectively, were introduced as *ansatz* in sub-section 4.4 of [1]. They have proved very useful and indispensable since their introduction. For instance, the new spacetime/intrinsic spacetime diagrams for the derivation of Lorentz transformation/intrinsic Lorentz transformation and their inverses in the four-world picture, (referred to as two-world picture in [1]), derived and presented as Figs. 8a and 8b of [1] (or Figs. 10a and 11a of part one of this paper [3]) and their inverses namely, Figs. 9a and 9b of [1], involve relative rotations of intrinsic affine spacetime coordinates, without any need for relative rotations of affine spacetime coordinates.

Once the intrinsic Lorentz transformation (ϕLT) and its inverse have been derived graphically as transformation of the primed intrinsic affine spacetime coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ of the intrinsic particle's frame into the unprimed intrinsic affine spacetime coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ of the intrinsic observer's frame and its inverse, then Lorentz transformation (LT) and its inverse in terms of primed affine spacetime coordinates $\tilde{x}', \tilde{y}', \tilde{z}'$ and $c\tilde{t}'$ of the particle's frame and the unprimed affine spacetime coordinates $\tilde{x}, \tilde{y}, \tilde{z}$ and $c\tilde{t}$ of the observer's frame can be written straight away, as the outward manifestations on flat four-dimensional spacetime of the intrinsic Lorentz transformation (ϕLT) and its inverse on flat two-dimensional intrinsic spacetime, as demonstrated in sub-section 4.4 of [1].

The indispensability of the flat two-dimensional proper intrinsic metric spacetime $(\phi\rho', \phi c\phi t')$ underlying flat four-dimensional proper metric spacetime (Σ', ct'), arises from the fact that it is possible for the intrinsic affine spacetime coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ of the intrinsic particle's frame $(\phi\tilde{x}', \phi c\phi\tilde{t}')$ that contains the one-dimensional intrinsic rest mass ϕm_0 of the particle in the intrinsic affine space coordinate $\phi\tilde{x}'$, to rotate anti-clockwise by an intrinsic angle $\phi\psi$ relative to the horizontal and vertical respectively and thereby project the intrinsic affine spacetime coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ of the intrinsic observer's frame $(\phi\tilde{x}, \phi c\phi\tilde{t})$ along the horizontal and vertical respectively, where the projective intrinsic affine space coordinate $\phi\tilde{x}$ of the observer's frame along the horizontal contains the one-dimensional intrinsic relativistic mass, $\phi m = \gamma\phi m_0$, of the particle, as happens in the first and second quadrants in Fig. 8a of [1], although the intrinsic rest mass ϕm_0 in the inclined $\phi\tilde{x}'$ and intrinsic relativistic mass ϕm in the projective $\phi\tilde{x}$ along the horizontal are not shown in that diagram.

The projective unprimed intrinsic affine coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ that constitute the observer's intrinsic frame, containing one-dimensional intrinsic relativistic mass ϕm of the particle in $\phi\tilde{x}$, are then made manifest outwardly in the unprimed affine spacetime coordinates $\tilde{x}$, $\tilde{y}$, $\tilde{z}$ and $c\tilde{t}$ of the observer's fame on flat four-dimensional spacetime, containing the three-dimensional relativistic mass, $m = \gamma m_0$, of the particle in affine 3-space $\tilde{\Sigma}(\tilde{x}, \tilde{y}, \tilde{z})$ of the observer's frame.

On the other hand, diagrams obtained by replacing the

inclined primed intrinsic affine coordinates $\phi\tilde{x}', \phi c\phi\tilde{t}', -\phi\tilde{x}'^*$ and $-\phi c\phi\tilde{t}'^*$ of the symmetry-partner intrinsic particles' frames $(\phi\tilde{x}', \phi c\phi\tilde{t}')$ and $(-\phi\tilde{x}'^*, -\phi c\phi\tilde{t}'^*)$ by inclined primed affine spacetime coordinates $\tilde{x}', c\tilde{t}', -\tilde{x}'^*$ and $-c\tilde{t}'^*$ respectively of the symmetry-partner particles' frames $(\tilde{x}', \tilde{y}', \tilde{z}', c\tilde{t}')$ and $(-\tilde{x}'^*, -\tilde{y}'^*, -\tilde{z}'^*, -c\tilde{t}'^*)$ in the positive and negative universes in Figs. 8a and 8b of [1], that is, by letting $\phi\tilde{x}' \to \tilde{x}'$; $\phi c\phi\tilde{t}' \to c\tilde{t}'$; $-\phi\tilde{x}'^* \to -\tilde{x}'^*$; $-\phi c\phi\tilde{t}'^* \to -c\tilde{t}'^*$; $\phi\tilde{x} \to \tilde{x}$; $\phi c\phi\tilde{t} \to c\tilde{t}$; $-\phi\tilde{x}^* \to -\tilde{x}^*$ and $-\phi c\phi\tilde{t}^* \to -c\tilde{t}^*$ in those diagrams, as would be done in the four-world picture in the absence of the intrinsic spacetime coordinates, are invalid or will not work.

The end of the foregoing paragraph is so since the affine space coordinates $\tilde{y}'$ and $\tilde{z}'$ of the particle's frame are not rotated along with the affine space coordinate $\tilde{x}'$ from affine 3-space $\tilde{\Sigma}'(\tilde{x}', \tilde{y}', \tilde{z}')$ of the particle's frame (as a hyper-surface) along the horizontal towards the time dimension $c\tilde{t}$ along the vertical. And the only rotated coordinate $\tilde{x}'$, which is inclined at angle ψ to the horizontal, cannot contain the three-dimensional rest mass m_0 of the particle, which can then be "projected" as three-dimensional relativistic mass, $m = \gamma m_0$, into the projective affine 3-space $\tilde{\Sigma}(\tilde{x}, \tilde{y}, \tilde{z})$ of the observer's frame (as a hyper-surface) along the horizontal. It then follows that the observational fact of the evolution of the rest mass m_0 of the particle into relativistic mass, $m = \gamma m_0$, in SR, is impossible in the context of diagrams involving rotations of the affine spacetime coordinates $\tilde{x}'$ and $c\tilde{t}'$ of the particle's frame relative to the affine spacetime coordinates $\tilde{x}$ and $c\tilde{t}$ of the observer's frame, which are in relative motion along their collinear $\tilde{x}'-$ and $\tilde{x}-$axes in the four-world picture. This rules out the possibility (or validity) of such diagrams in the four-world picture. As noted in [1], if such diagrams are drawn, it must be understood that they are hypothetical or intrinsic (i.e. non-observable).

Further more, it is possible for the intrinsic affine spacetime coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ of the particle's intrinsic frame $(\phi\tilde{x}', \phi c\phi\tilde{t}')$, containing the one-dimensional intrinsic rest mass ϕm_0 of the particle in the intrinsic affine space coordinate $\phi\tilde{x}'$, to rotate relative to their projective affine intrinsic spacetime coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ of the observer's intrinsic frame $(\phi\tilde{x}, \phi c\phi\tilde{t})$, that contains the 'projective' one-dimensional intrinsic relativistic mass, $\phi m = \gamma\phi m_0$, of the particle in the projective intrinsic affine space coordinate $\phi\tilde{x}$, by intrinsic angles $\phi\psi$ larger that $\frac{\pi}{2}$, that is, in the range $\frac{\pi}{2} < \phi\psi \leq \pi$, (assuming rotation by $\phi\psi = \frac{\pi}{2}$ can be avoided), in Fig. 8a of [1]. This will make the particle's intrinsic frame $(\phi\tilde{x}', \phi c\phi\tilde{t}')$ containing the positive intrinsic rest mass ϕm_0 of the particle in the inclined affine intrinsic coordinate $\phi\tilde{x}'$ in the positive universe to make transition into the negative universe through the second quadrant to become particle's intrinsic frame $(-\phi\tilde{x}'^*, -\phi c\phi\tilde{t}'^*)$ containing negative intrinsic rest mass $-\phi m_0^*$ of the particle in the negative intrinsic affine space coordinate $-\phi\tilde{x}'^*$, as explained in section 2 of [2].

The negative intrinsic affine spacetime coordinates $-\phi\tilde{x}'^*$ and $-\phi c\phi\tilde{t}'^*$ of the intrinsic particle's frame, into which the positive intrinsic affine coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ of the particle's intrinsic frame in the positive universe transform upon making transition into the negative universe through the second quadrant, will be inclined intrinsic affine coordinates in the second quadrant and the third quadrant respectively. They will project intrinsic affine coordinates $-\phi\tilde{x}^*$ and $-\phi c\phi\tilde{t}^*$ of the observer's intrinsic frame along the horizontal and vertical respectively in the third quadrant. Thus the observer's intrinsic frame $(-\phi\tilde{x}^*, -\phi c\phi\tilde{t}^*)$ containing negative intrinsic relativistic mass, $-\phi m^* = -\gamma\phi m_0^*$, in the intrinsic affine space coordinate $-\phi\tilde{x}^*$, will automatically appear in the negative universe, upon the particle's intrinsic frame $(\phi\tilde{x}', \phi c\phi\tilde{t}')$ containing positive intrinsic rest mass ϕm_0 of the particle in the first quadrant making transition into the second quadrant. The observer's intrinsic frame $(-\phi\tilde{x}^*, -\phi c\phi\tilde{t}^*)$ containing relativistic intrinsic mass $-\phi m^* = -\gamma\phi m_0^*$ in $-\phi x^*$ will then be made manifest in observer's frame $(-\tilde{x}^*, -\tilde{y}^*, -\tilde{z}^*, -c\tilde{t}^*)$ on flat spacetime of the negative universe, containing negative three-dimensional relativistic mass, $-m^* = -\gamma m_0^*$, of the particle.

It is therefore possible for a particle in relative motion in the positive universe to make transition into the negative universe in the context of the geometrical representation of ϕLT/LT in the two-world picture (now re-identified as four-world picture) in Figs. 8a and 8b of [1], assuming rotation of intrinsic affine spacetime coordinates $\phi\tilde{x}'$ and $\phi c\phi\tilde{t}'$ of the particle's intrinsic frame relative to the intrinsic affine spacetime coordinates $\phi\tilde{x}$ and $\phi c\phi\tilde{t}$ of the observer's intrinsic frame by intrinsic angle $\phi\psi = \frac{\pi}{2}$, corresponding to intrinsic speed $\phi v = \phi c$ of relative intrinsic motion, can be avoided in the process of rotation by $\phi\psi > \frac{\pi}{2}$.

On the other hand, letting the affine spacetime coordinates $\tilde{x}'$ and $c\tilde{t}'$ of the particle's frame $(\tilde{x}', \tilde{y}', \tilde{z}', c\tilde{t}')$ to rotate relative to the affine spacetime coordinates $\tilde{x}$ and $c\tilde{t}$ respectively of the observer's frame $(\tilde{x}, \tilde{y}, \tilde{z}, c\tilde{t})$ in the positive universe by angle ψ larger than $\frac{\pi}{2}$, that is in the range $\frac{\pi}{2} < \psi \leq \pi$, (assuming $\psi = \frac{\pi}{2}$ can be avoided), will cause the affine spacetime coordinates $\tilde{x}'$ and $c\tilde{t}'$ to make transition into the negative universe through the second quadrant to become inclined affine coordinates $-\tilde{x}'^*$ and $-c\tilde{t}'^*$ in the second and third quadrants respectively. However the non-rotated affine space coordinates $\tilde{y}'$ and $\tilde{z}'$ of the particle's frame will remain along the horizontal in the first quadrant in the positive universe. This situation in which only two of four coordinates of a frame make transition from the positive universe into the negative universe is impossible.

Moreover since the three-dimensional rest mass m_0 of the particle cannot be contained in the only rotated affine space coordinate $\tilde{x}'$, the rest mass of the particle will be unable to make transition into the negative universe with the rotated coordinates $\tilde{x}'$ and $c\tilde{t}'$. It is therefore impossible for a particle in relative motion in the positive universe to make transition into the negative universe in the context of diagrams involving rotation of affine spacetime coordinates $\tilde{x}'$ and $c\tilde{t}'$ of the particle's frame relative to affine spacetime coordinates $\tilde{x}$ and

$c\tilde{t}$ of the observer's frame, where the two frames are in motion along their collinear $\tilde{x}'-$ and $\tilde{x}-$axes, in the two-world picture (now re-identified as four-world picture). This further renders such diagrams ineffective and impossible.

Relative rotations of intrinsic affine spacetime coordinates in the spacetime/intrinsic spacetime diagrams for deriving intrinsic Lorentz transformation/Lorentz transformation are unavoidable in the present many-world picture. This makes the flat two-dimensional intrinsic metric spacetime underlying flat four-dimensional metric spacetime indispensable in the context of the present theory.

1.2 Origin of the intrinsic space and intrinsic time dimensions

It has been shown that the quartet of Euclidean 3-spaces and underlying one-dimensional intrinsic spaces in Fig. 2 of part one of this paper [3], simplifies naturally as Figs. 6a and 6b of that paper, where Fig. 6a is valid with respect to 3-observers in our proper Euclidean 3-space Σ' and 3-observer* in the proper Euclidean 3-space $-\Sigma'^{*}$ of the negative universe and Fig. 6b is valid with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe and 3-observer* in the proper Euclidean 3-space $-\Sigma^{0\prime *}$ of the negative time-universe, as indicated in those diagrams. Figs. 6a and 6b of [3] ultimately transform into Figs. 8a and 8b respectively of that paper naturally with respect to the same 3-observers in the proper Euclidean 3-spaces with respect to whom Figs. 6a and 6b are valid.

The one-dimensional proper intrinsic spaces underlying the proper Euclidean 3-spaces have been introduced without deriving them in the first part of this paper [3]. Now let us assume that the underlying one-dimensional proper intrinsic spaces have not been known in Figs. 6a and 6b of [3]. Then let us reproduce the first quadrant of those figures without the intrinsic spaces as Figs. 1a and 1b respectively here.

The one-dimensional proper (or classical) space $\rho^{0\prime}$ along the vertical in Fig. 1a (to which the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe naturally contracts with respect to 3-observers in our proper Euclidean 3-space Σ'), projects a component to be denoted by ρ' into our proper Euclidean 3-space Σ' (considered as a hyper-surface along the horizontal), which is given as follows:

$$\rho' = \rho^{0\prime} \cos\psi_0 = \rho^{0\prime} \cos\frac{\pi}{2} = 0 \qquad (1)$$

where the fact that $\rho^{0\prime}$ is naturally inclined at absolute angle $\psi_0 = \frac{\pi}{2}$ to the horizontal, corresponding to absolute speed $V_0 = c$ of every point along $\rho^{0\prime}$ relative to 3-observers in Σ' (discussed extensively in sub-section 1.1 of [3]) has been used in (1).

Equation (1) states that the one-dimensional space $\rho^{0\prime}$ along the vertical projects zero component (or nothing) into the Euclidean 3-space Σ' (as a hyper-surface) along the horizontal. However we shall not ascribe absolute nothingness

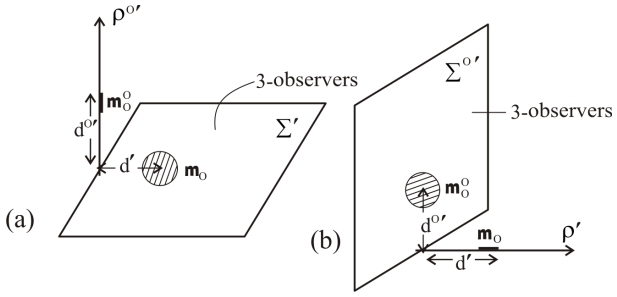


Fig. 1: (a) The proper Euclidean 3-space of our universe Σ' (as a hyper-surface along the horizontal), containing the rest mass m_0 of an object and the one-dimensional proper space $\rho^{0\prime}$ containing the one-dimensional rest mass m_0^0 of the symmetry-partner object in the positive time-universe relative to 3-observers in Σ'; $\rho^{0\prime}$ containing one-dimensional m_0^0 being the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe containing three-dimensional rest mass m_0^0 with respect to 3-observers in $\Sigma^{0\prime}$. (b) The proper Euclidean 3-space of the positive time-universe $\Sigma^{0\prime}$ (as a hyper-surface along the vertical), containing the rest mass m_0^0 of an object and the one-dimensional proper space ρ' containing the one-dimensional rest mass m_0 of the symmetry-partner object in our universe relative to 3-observers in $\Sigma^{0\prime}$; ρ' containing one-dimensional m_0 being the proper Euclidean 3-space Σ' of the positive (or our) universe containing three-dimensional rest mass m_0 with respect to 3-observers in Σ'.

to the projection of the physical one-dimensional space $\rho^{0\prime}$ along the vertical into the Euclidean 3-space Σ' along the horizontal in Fig. 1a. The one-dimensional space $\rho^{0\prime}$ certainly "casts a shadow" into Σ'.

Actually, it is the factor $\cos\frac{\pi}{2}$ that vanishes in (1) and not $\rho^{0\prime}$ multiplying it. Thus let us re-write (1) as follows:

$$\rho' = \rho^{0\prime} \cos\frac{\pi}{2} = 0 \times \rho^{0\prime} \equiv \phi\rho' \qquad (2)$$

where $\phi\rho'$ is without the superscript "0" label because it lies in (or underneath) our Euclidean 3-space Σ' (without superscript "0" label) along the horizontal.

Thus instead of associating absolute nothingness to the projection of $\rho^{0\prime}$ along the vertical into the Euclidean 3-space Σ' along the horizontal, as done in (1), a dimension $\phi\rho'$ of intrinsic (that is, non-observable and non-detectable) quality, has been attributed to it in (2). Hence $\phi\rho'$ shall be referred to as intrinsic space. It is proper (or classical) intrinsic space by virtue of its prime label.

Any interval of the one-dimensional intrinsic space (or intrinsic space dimension) $\phi\rho'$ is equivalent to zero interval of the one-dimensional physical space $\rho^{0\prime}$, (as follows from $\phi\rho' \equiv 0 \times \rho^{0\prime}$ in (2)). It then follows that any interval of the proper intrinsic space $\phi\rho'$ is equivalent to zero distance of the physical proper Euclidean 3-space Σ'. Or any interval of $\phi\rho'$ is no interval of space. The name nospace shall be coined for $\phi\rho'$ from the last statement, as an alternative to intrinsic space, where $\phi\rho'$ is proper (or classical) nospace by virtue of

the prime label on it.

As derived in sub-section 1.2 of the first part of this paper [3], the Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe is geometrically contracted to the one-dimensional space $\rho^{0\prime}$ with respect to 3-observers in our Euclidean 3-space Σ' between Fig. 3 and Fig. 6a of [3], where $\rho^{0\prime}$ can be considered to be along any direction of the Euclidean 3-space $\Sigma^{0\prime}$ that contracts to it, with respect to 3-observers in Σ'. Thus $\rho^{0\prime}$ is an isotropic one-dimensional space with no unique orientation in the Euclidean 3-space $\Sigma^{0\prime}$ that contracts to it with respect to 3-observers in Σ'. The one-dimensional intrinsic space (or one-dimensional nospace) $\phi\rho'$, which $\rho^{0\prime}$ projects into the Euclidean 3-space Σ', is consequently an isotropic intrinsic space dimension with no unique orientation in Σ' with respect to 3-observers in Σ'.

The one-dimensional proper (or classical) space ρ' along the horizontal in Fig. 1b, to which our proper Euclidean 3-space Σ' geometrically contracts with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe, as explained between Fig. 4 and Fig. 6b in sub-section 1.2 of [3], likewise projects one-dimensional proper intrinsic space (or proper nospace) $\phi\rho^{0\prime}$ into the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe along the vertical in Fig. 1b (not yet shown in Fig. 1b), where $\phi\rho^{0\prime}$ is an isotropic one-dimensional intrinsic space dimension (with no unique orientation) in $\Sigma^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$.

As follows from all the foregoing, Fig. 1a must be replaced with Fig. 2a, where the one-dimensional proper intrinsic space $\phi\rho'$ projected into the proper Euclidean 3-space Σ' by the one-dimensional proper space $\rho^{0\prime}$ with respect to 3-observers in Σ' has been shown. Fig. 1b must likewise be replaced with Fig 2b, where the one-dimensional proper intrinsic space (or proper nospace) $\phi\rho^{0\prime}$ projected into the proper Euclidean 3-space $\Sigma^{0\prime}$ by the one-dimensional proper space ρ' with respect to 3-observers in $\Sigma^{0\prime}$ has been shown.

The one-dimensional isotropic proper (or classical) intrinsic space $\phi\rho'$ underlying the proper (or classical) Euclidean 3-space Σ' of the positive (or our) universe with respect to 3-observers in Σ' and the one-dimensional isotropic proper (or classical) intrinsic space $\phi\rho^{0\prime}$ underlying the proper (or classical) Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$, have thus been derived. The derivations of the proper intrinsic space $-\phi\rho'^{*}$ underlying the proper Euclidean 3-space $-\Sigma'^{*}$ of the negative universe with respect to 3-observers in $-\Sigma'^{*}$ and of $-\phi\rho^{0\prime *}$ underlying the proper Euclidean 3-space $-\Sigma^{0\prime *}$ of the negative time-universe with respect to 3-observers* in $-\Sigma'^{0*}$, follow directly from the derivations of $\phi\rho'$ underlying Σ' and $\phi\rho^{0\prime}$ underlying $\Sigma^{0\prime}$ above.

Following the introduction of the flat 2-dimensional proper intrinsic spacetimes $(\phi\rho', \phi c\phi t')$ and $(-\phi\rho'^{*}, -\phi c\phi t'^{*})$ that underlie the flat four-dimensional proper spacetimes (Σ', ct') and $(-\Sigma'^{*}, -ct'^{*})$ of the positive (or our) universe and the negative universe respectively as *ansatz* in sub-section 4.4 of [1],

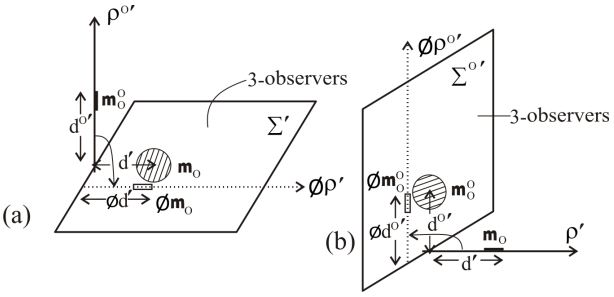


Fig. 2: (a) The one-dimensional proper space $\rho^{0\prime}$ containing one-dimensional rest mass m_0^0 along the vertical, projects one-dimensional proper intrinsic space $\phi\rho'$ containing one-dimensional intrinsic rest mass ϕm_0 into the proper Euclidean 3-space Σ' (as a hyper-surface) containing the rest mass m_0 along the horizontal, with respect to 3-observers in Σ'. (b) The one-dimensional proper space ρ' containing one-dimensional rest mass m_0 along the horizontal, projects one-dimensional proper intrinsic space $\phi\rho^{0\prime}$ containing one-dimensional intrinsic rest mass ϕm_0^0 into the proper Euclidean 3-space $\Sigma^{0\prime}$ (as a hyper-surface) containing rest mass m_0^0 along the vertical, with respect to 3-observers in $\Sigma^{0\prime}$

the one-dimensional proper intrinsic spaces $\phi\rho'$ and $-\phi\rho'^{*}$ underlying the proper Euclidean 3-spaces Σ' and $-\Sigma'^{*}$ of the positive and negative universes and the proper intrinsic spaces $\phi\rho^{0\prime}$ and $-\phi\rho^{0\prime *}$ underlying the proper Euclidean 3-spaces $\Sigma^{0\prime}$ and $-\Sigma^{0\prime *}$ of the positive and negative time-universes were introduced without deriving them in Figs. 2, 3 and 4 and Figs. 6a and 6b of the first part of this paper [3]. The existence in nature of the one-dimensional isotropic intrinsic spaces underlying the physical Euclidean 3-spaces has now been validated.

1.3 Origin of one-dimensional intrinsic rest mass in one-dimensional proper intrinsic space underlying rest mass in proper Euclidean 3-space

The one-dimensional proper space $\rho^{0\prime}$, being orthogonal to the proper Euclidean 3-space Σ' (as a hyper-surface) along the horizontal, possesses absolute speed $V_0 = c$ at every point along its length with respect to 3-observers in Σ', as has been well discussed in sub-section 1.1 of [3]. Consequently, the one-dimensional rest mass m_0^0 of a particle or object in $\rho^{0\prime}$ acquires the absolute speed $V_0 = c$ of $\rho^{0\prime}$ with respect to 3-observers in Σ' in Figs. 1a and 2a.

On the other hand, the Euclidean 3-space Σ' being along the horizontal (as a hyper-surface), possesses zero absolute speed ($V_0 = 0$) at every point of it with respect to 3-observers in Σ'. The projective intrinsic space (or nospace) $\phi\rho'$, being along the horizontal, likewise possesses zero absolute intrinsic speed ($\phi V_0 = 0$) at every point along its length with respect to 3-observers in Σ' in Fig. 2a.

The one-dimensional rest mass m_0^0 in the one-dimensional proper space $\rho^{0\prime}$ along the vertical in Figs. 1a or 2a, can be said to be in non-detectable absolute motion at constant

absolute speed $V_0 = c$ along $\rho^{0\prime}$ with respect to 3-observers in the proper Euclidean 3-space Σ' in that figure. There is a mass relation in the context of absolute motion that can be applied for the non-detectable absolute motion at absolute speed $V_0 = c$ of m_0^0 along $\rho^{0\prime}$, which shall be derived elsewhere in the systematic development of the present theory. It shall be temporarily written hereunder because of the need to use it at this point.

Let us revisit Fig. 7 of part one of this paper [3], drawn to illustrate the concept of time and intrinsic time induction only. It is assumed that the proper intrinsic metric space $\phi\rho^{0\prime}$ possesses absolute intrinsic speed $\phi V_0 < \phi c$ at every point along its length, thereby causing $\phi\rho^{0\prime}$ to be inclined at a constant absolute intrinsic angle, $\phi\psi_0 < \frac{\pi}{2}$, relative to its projection $\phi\rho'$ along the horizontal in that figure. This is so since the uniform absolute intrinsic speed ϕV_0 along the length of $\phi\rho^{0\prime}$ is related to the constant absolute intrinsic angle $\phi\psi_0$ of inclination to the horizontal of $\phi\rho^{0\prime}$ as, $\sin\phi\psi_0 = \phi V_0/\phi c$, (see Eq. (1) of [3]). It follows from this relation that when the inclined $\phi\rho^{0\prime}$ lies along the horizontal, thereby being the same as its projection $\phi\rho'$ along the horizontal, it possesses constant zero absolute intrinsic speed ($\phi V_0 = 0$) at every point along its length along the horizontal with respect to the 3-observer in Σ' in that figure, just as it has been said that the projective $\phi\rho'$ along the horizontal possesses absolute intrinsic speed $V_0 = 0$ at every point along its length with respect to 3-observers Σ' in Fig. 2a earlier. And for $\phi\rho^{0\prime}$ to lie along the vertical in Fig. 7 of [3], it possesses constant absolute intrinsic speed $\phi V_0 = \phi c$ at every point along its length with respect to the 3-observer in Σ'.

Now let a one-dimensional intrinsic rest mass ϕm_0^0 be located at any point along the inclined proper intrinsic metric space $\phi\rho^{0\prime}$ in Fig. 7 of [3]. Then ϕm_0^0 will acquire absolute intrinsic speed $\phi V_0 < \phi c$ along the inclined $\phi\rho^{0\prime}$. It will project another intrinsic rest mass ϕm_0 (since it is not in relative motion) into the proper intrinsic space $\phi\rho'$, which the inclined $\phi\rho^{0\prime}$ projects along the horizontal. The relation between the 'projective' intrinsic rest mass ϕm_0 in the projective proper intrinsic space $\phi\rho'$ along the horizontal and the intrinsic rest mass ϕm_0^0 along the inclined proper intrinsic space $\phi\rho^{0\prime}$ (not shown in Fig. 7 of [3]), is the intrinsic mass relation in the context of absolute intrinsic motion to be derived formally elsewhere. It is given as follows:

$$\phi m_0 = \phi m_0^0 \cos^2\phi\psi_0 = \phi m_0^0\left(1 - \frac{\phi V_0^2}{\phi c^2}\right) \tag{3}$$

The outward manifestation in the proper 3-dimensional Euclidean space Σ' (in Fig. 7 of [3]) of Eq. (3), obtained by simply removing the symbol ϕ, is the following

$$m_0 = m_0^0 \cos^2\psi_0 = m_0^0\left(1 - \frac{V_0^2}{c^2}\right) \tag{4}$$

Corresponding to relations (3) and (4) in the contexts of absolute intrinsic motion and absolute motion, there are the intrinsic mass relation in the context of relative intrinsic motion (or in the context of intrinsic special theory of relativity (ϕSR)) and mass relation in the context of relative motion (or in the context of SR). The generalized forms involving intrinsic angle $\phi\psi$ and angle ψ of intrinsic mass relation in the context of ϕSR and mass relation in the context of SR, derived and presented as Eqs. (15) and (16) in section 3 of [2] are the following

$$\phi m = \phi m_0 \sec\phi\psi = \phi m_0\left(1 - \frac{\phi v^2}{\phi c^2}\right)^{-1/2} \tag{5}$$

and

$$m = m_0 \sec\psi = m_0\left(1 - \frac{v^2}{c^2}\right)^{-1/2} \tag{6}$$

One finds that relations (3) and (4) in the context of absolute intrinsic motion and absolute motion differ grossly from the corresponding relations (5) and (6) in relative intrinsic motion (or in the context of ϕSR) and in relative motion (or in the context of SR).

Since the one-dimensional rest mass m_0^0 possesses absolute speed $V_0 = c$ of non-detectable absolute motion along $\rho^{0\prime}$ with respect to 3-observers in Σ' in Fig. 2a, relation (4) can be applied for the "projection" of m_0^0 into the Euclidean 3-space Σ' with respect to 3-observers in Σ' in that figure. We must simply let $\psi_0 = \frac{\pi}{2}$ and $V_0 = c$ in Eq. (4) to have

$$m_0 = m_0^0 \cos^2\frac{\pi}{2} = m_0^0\left(1 - \frac{c^2}{c^2}\right) = 0 \tag{7}$$

Equation (7) states that the one-dimensional rest mass m_0^0 in the one-dimensional space $\rho^{0\prime}$ along the vertical in Fig. 2a, (to which the three-dimensional rest mass m_0^0 in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$ contracts relative to 3-observers in our Euclidean 3-space Σ'), projects zero rest mass (or nothing) into our Euclidean 3-space Σ' along the horizontal. However the one-dimensional rest mass m_0^0 in $\rho^{0\prime}$ along the vertical certainly 'casts a shadow' into the Euclidean 3-space Σ' considered as a hyper-surface along the horizontal in Fig. 2a.

It is the factor $\cos^2\frac{\pi}{2}$ or $(1 - c^2/c^2)$ that vanishes and not the rest mass m_0^0 multiplying it in Eq. (7). Thus let us re-write Eq. (7) as follows:

$$m_0 = m_0^0 \cos^2\frac{\pi}{2} = 0 \times m_0^0 \equiv \phi m_0 \tag{8}$$

Instead of ascribing absolute nothingness to the "projection" of the one-dimensional rest mass m_0^0 in the one-dimensional space $\rho^{0\prime}$ along the vertical into our proper Euclidean 3-space as a hyper-surface Σ' along the horizontal in Fig. 2a in Eq. (7), a one-dimensional quantity ϕm_0 of intrinsic (that is, nonobservable and non-detectable) quality has been ascribed to it in Eq. (8). Hence ϕm_0 shall be referred to as intrinsic rest mass.

Any quantity of the one-dimensional intrinsic rest mass ϕm_0 is equivalent to zero quantity of the one-dimensional rest

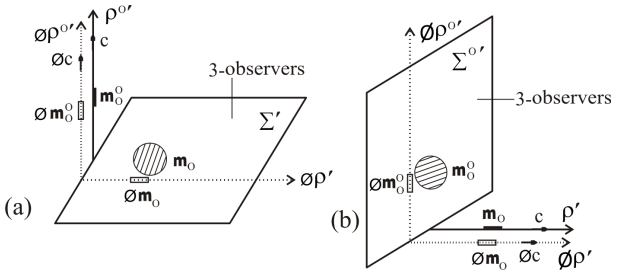


Fig. 3: (a) The proper intrinsic space $\phi\rho^{0\prime}$ containing intrinsic rest mass ϕm_0^0, projected into the proper Euclidean 3-space $\Sigma^{0\prime}$ along the vertical by one-dimensional proper space ρ' containing one-dimensional rest mass m_0 along the horizontal in Fig. 2b, is added to Fig. 2a, where it lies parallel to $\rho^{0\prime}$ along the vertical, giving rise to a flat four-dimensional proper space $(\Sigma', \rho^{0\prime})$ underlied by flat two-dimensional proper intrinsic space $(\phi\rho', \phi\rho^{0\prime})$ with respect to 3-observers in Σ'. (b) The proper intrinsic space $\phi\rho'$ containing intrinsic rest mass ϕm_0, projected into the proper Euclidean 3-space Σ' along the horizontal by one-dimensional proper space $\rho^{0\prime}$ containing one-dimensional rest mass m_0^0 along the vertical in Fig. 2a, is added to Fig. 2b, where it lies parallel to ρ' along the horizontal, giving rise to a flat four-dimensional proper space $(\Sigma^{0\prime}, \rho')$ underlied by flat two-dimensional proper intrinsic space $(\phi\rho^{0\prime}, \phi\rho')$ with respect to 3-observers in $\Sigma^{0\prime}$.

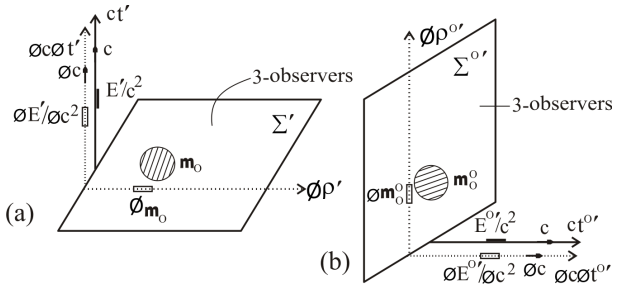


Fig. 4: (a) The one-dimensional proper space $\rho^{0\prime}$ and the proper intrinsic space $\phi\rho^{0\prime}$ along the vertical with respect to 3-observers in Σ' in Fig. 3a, transform into the proper time dimension ct' and proper intrinsic time dimension $\phi c\phi t'$ respectively, giving rise to a flat four-dimensional proper spacetime (Σ', ct') underlied by flat two-dimensional proper intrinsic spacetime $(\phi\rho', \phi c\phi t')$ with respect to 3-observers in Σ'. (b) The one-dimensional proper space ρ' and the proper intrinsic space $\phi\rho'$ along the horizontal with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 3b, transform into the proper time dimension $ct^{0\prime}$ and proper intrinsic time dimension $\phi c\phi t^{0\prime}$ respectively, giving rise to a flat four-dimensional proper spacetime $(\Sigma^{0\prime}, ct^{0\prime})$ underlied by flat two-dimensional proper intrinsic spacetime $(\phi\rho^{0\prime}, \phi c\phi t^{0\prime})$ with respect to 3-observers in $\Sigma^{0\prime}$.

mass m_0^0 in the one-dimensional space $\rho^{0\prime}$, as follows from $\phi m_0 \equiv 0 \times m_0^0$ in Eq. (8). It then follows that any quantity of the intrinsic rest mass ϕm_0 is equivalent to zero quantity of three-dimensional rest mass m_0 in Σ'. Or any quantity of intrinsic rest mass is no rest mass. An alternative name coined from the preceding statement namely, nomass, shall be given to the intrinsic rest mass ϕm_0. The intrinsic rest mass ϕm_0 in the proper (or classical) intrinsic space is the proper (or classical) nomass.

The 'projective' intrinsic rest mass (or proper nomass) ϕm_0 in the projective proper intrinsic space $\phi\rho'$, lies directly underneath the rest mass m_0 in the proper Euclidean 3-space Σ', as already shown in Fig. 2a. The one-dimensional rest mass m_0 in the one-dimensional proper (or classical) space ρ' along the horizontal in Fig. 2b, likewise "projects" intrinsic rest mass (or proper nomass) ϕm_0^0 into the projective proper (or classical) intrinsic space $\phi\rho^{0\prime}$, which lies directly underneath the rest mass m_0^0 in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$, as already shown in Fig. 2b.

Now the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in it in Fig. 2b, is what appears as one-dimensional proper space $\rho^{0\prime}$ along the vertical with respect to 3-observers in our proper Euclidean 3-space Σ' in Fig. 2a. The one-dimensional proper intrinsic space $\phi\rho^{0\prime}$ projected into (or underneath) $\Sigma^{0\prime}$ by ρ' along the horizontal in Fig. 2b, must be added to Fig. 2a, where it must lie parallel to $\rho^{0\prime}$ along the vertical, thereby converting Fig. 2a to Fig. 3a with respect to 3-observers in Σ'. The one-dimensional proper intrinsic space $\phi\rho'$ projected into (or underneath) our proper Euclidean 3-space Σ' by $\rho^{0\prime}$ along the vertical in Fig. 2a, must likewise be added to Fig. 2b, where it must lie parallel to ρ' along the horizontal, thereby converting Fig. 2b to Fig. 3b with respect to 3-observers in $\Sigma^{0\prime}$.

Finally, as explained for the transformations of Figs. 6a and 6b into Figs. 8a and 8b respectively in sub-section 1.3 of part one of this paper [3], the one-dimensional proper (or classical) space $\rho^{0\prime}$ and the one-dimensional proper (or classical) intrinsic space $\phi\rho^{0\prime}$ lying parallel to it along the vertical in Fig. 3a, transform into the proper time dimension ct' and the proper intrinsic time dimension $\phi c\phi t'$ of the positive (or our) universe with respect to 3-observers in our proper Euclidean 3-space Σ', thereby converting Fig. 3a to the final Fig. 4a.

The one-dimensional proper (or classical) space ρ' and the one-dimensional proper (or classical) intrinsic space $\phi\rho'$ lying parallel to it along the horizontal in Fig. 3b, likewise transform into the proper time dimension $ct^{0\prime}$ and the proper intrinsic time dimension $\phi c\phi t^{0\prime}$ of the positive time-universe with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe, thereby converting Fig. 3b to the final Fig. 4b.

As also explained in drawing Figs. 9a and 9b of [3], the one-dimensional rest mass m_0^0 in the one-dimensional proper (or classical) space $\rho^{0\prime}$ and the one-dimensional intrinsic rest mass ϕm_0^0 in the proper (or classical) intrinsic space $\phi\rho^{0\prime}$ in Fig. 3a must be replaced by one-dimensional equivalent rest mass E'/c^2, where $E' = m_0^0 c^2$, in the proper time-dimension ct' and one-dimensional equivalent intrinsic rest mass $\phi E'/\phi c^2$, where $\phi E' = \phi m_0^0 \phi c^2$, in the proper intrinsic time

dimension $\phi c\phi t'$ respectively, as done in Fig. 4a. The one-dimensional rest mass m_0 in ρ' and the intrinsic rest mass ϕm_0 in $\phi\rho'$ along the horizontal in Fig. 3b must likewise be replaced by $E^{0\prime}/c^2$; $E^{0\prime} = m_0c^2$, in $ct^{0\prime}$ and $\phi E^{0\prime}/\phi c^2$; $\phi E^{0\prime} = \phi m_0\phi c^2$, in $\phi c\phi t^{0\prime}$ respectively, as done in Fig. 4b.

Fig. 4a now has flat two-dimensional proper intrinsic spacetime (or proper nospace-notime) $(\phi\rho', \phi c\phi t')$, containing intrinsic rest mass (or proper nomass) ϕm_0 (in $\phi\rho'$) and equivalent intrinsic rest mass $\phi E'/\phi c^2$ (in $\phi c\phi t'$), underlying flat four-dimensional proper spacetime (Σ', ct'), containing rest mass m_0 (in Σ') and equivalent rest mass E'/c^2 (in ct'), of the positive (or our) universe. Fig. 4b likewise now has flat two-dimensional proper intrinsic spacetime $(\phi\rho^{0\prime}, \phi c\phi t^{0\prime})$, containing intrinsic rest mass ϕm_0^0 (in $\phi\rho^{0\prime}$) and equivalent intrinsic rest mass $\phi E^{0\prime}/\phi c^2$ (in $\phi c\phi t^{0\prime}$), underlying flat proper spacetime $(\Sigma^{0\prime}, ct^{0\prime})$, containing rest mass m_0^0 (in $\Sigma^{0\prime}$) and equivalent rest mass $E^{0\prime}/c^2$ (in $ct^{0\prime}$), of the positive time-universe.

In tracing the origin of the proper intrinsic space $\phi\rho'$ and the intrinsic rest mass ϕm_0 contained in it in Fig. 4a, we find that the one-dimensional proper space $\rho^{0\prime}$ containing one-dimensional rest mass m_0^0 along the vertical with respect to 3-observers in the proper Euclidean 3-space Σ' of the positive (or our) universe in Fig. 2a or 3a, projects proper intrinsic space $\phi\rho'$ containing intrinsic rest mass ϕm_0 into the proper Euclidean 3-space Σ'. Then as $\rho^{0\prime}$ containing m_0^0 along the vertical in Fig. 2a or 3a (being along the vertical) naturally transforms into proper time dimension ct' containing equivalent rest mass E'/c^2 with respect to 3-observers in Σ' in Fig. 4a, its projection $\phi\rho'$ containing ϕm_0 into Σ' along the horizontal (being along the horizontal) remains unchanged with respect to 3-observers in Σ'.

The conclusion then is that the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$, (which is one-dimensional space $\rho^{0\prime}$ with respect to 3-observers in our proper Euclidean 3-space Σ'), is ultimately the origin of the one-dimensional proper intrinsic space $\phi\rho'$ underlying the proper Euclidean 3-space Σ' of our universe and the three-dimensional rest mass m_0^0 of a particle of object in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$ is the origin of the one-dimensional intrinsic rest mass ϕm_0 in the proper intrinsic space $\phi\rho'$ lying directly underneath the rest mass m_0 of the symmetry-partner particle or object in the proper Euclidean 3-space Σ' of our universe. In other words, the proper Euclidean 3-space $\Sigma^{0\prime}$ containing the three-dimensional rest mass m_0^0 of a particle or object in the positive time-universe, "casts a shadow" of one-dimensional isotropic proper intrinsic space $\phi\rho'$ containing one-dimensional intrinsic rest mass ϕm_0 into the proper Euclidean 3-space Σ' containing the rest mass m_0 of the symmetry-partner particle or object in our universe, where ϕm_0 in $\phi\rho'$ lies directly underneath m_0 in Σ'.

And in tracing the origin of the proper intrinsic time dimension $\phi c\phi t'$ that contains the equivalent intrinsic rest mass $\phi E'/\phi c^2$, lying parallel to the proper time dimension ct' containing the equivalent rest mass E'/c^2 in Fig. 4a, we find that the one-dimensional proper space ρ' containing one-dimensional rest mass m_0 along the horizontal with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe in Fig. 2b or 3b, where ρ' is the Euclidean 3-space Σ' of our universe with respect to 3-observers in Σ', as derived between Fig. 4 and Fig. 6b in sub-section 1.2 of [3], projects one-dimensional proper intrinsic space $\phi\rho^{0\prime}$ containing intrinsic rest mass ϕm_0^0 underneath the proper Euclidean 3-space $\Sigma^{0\prime}$ containing rest mass m_0^0 of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 2b or 3b. The proper Euclidean 3-space $\Sigma^{0\prime}$ containing the rest mass m_0^0 and its underlying proper intrinsic space $\phi\rho^{0\prime}$ containing intrinsic rest mass ϕm_0^0 with respect to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe in Fig. 3b, are the proper time dimension ct' of our universe containing equivalent rest mass E'/c^2 and its underlying proper intrinsic time dimension $\phi c\phi t'$ of our universe containing equivalent intrinsic rest mass $\phi E'/\phi c^2$ with respect to 3-observers in Σ' in Fig. 4a.

The conclusion then is that the proper Euclidean 3-space Σ' of the positive (or our) universe is the origin of the proper intrinsic time dimension $\phi c\phi t'$ that lies parallel to the proper time dimension ct' of the positive (or our) universe in Fig. 4a. The three-dimensional rest mass m_0 of a particle or object in the proper Euclidean 3-space Σ' of our universe is the origin of the one-dimensional equivalent intrinsic rest mass $\phi E'/\phi c^2$ in the proper intrinsic time dimension $\phi c\phi t'$ that lies besides the one-dimensional equivalent rest mass E'/c^2 in the proper time dimension ct' of our universe in Fig. 4a.

The two-dimensional proper intrinsic metric spacetime (or proper metric nospace-notime) $(\phi\rho', \phi c\phi t')$, containing intrinsic rest mass ϕm_0 in $\phi\rho'$ and equivalent intrinsic rest mass $\phi E/\phi c^2$ in $\phi c\phi t'$, which underlies the flat proper metric spacetime (Σ', ct'), containing rest mass m_0 in Σ' and equivalent rest mass E'/c^2 in ct', has thus been derived within the four-world picture. The intrinsic special theory of relativity (ϕSR) operates on the flat proper intrinsic metric spacetime $(\phi\rho', \phi c\phi t')$ and the special theory of relativity (SR) operates on the flat proper metric spacetime (Σ', ct') in the absence of relativistic gravitational field. The flat two-dimensional proper intrinsic spacetime was introduced as *ansatz* in section 4.4 of [1] and it has proved indispensable in the present theory since then, as discussed fully earlier in sub-section 1.1 of this paper.

The derivations of the flat two-dimensional proper intrinsic spacetime $(\phi\rho', \phi c\phi t')$ containing intrinsic rest masses $(\phi m_0, \phi E'/\phi c^2)$ of particles and bodies, which underlies the flat four-dimensional proper spacetime (Σ', ct') containing the rest masses $(m_0, E'/c^2)$ of particles and bodies in our universe and in the other three universes, as presented in this sub-section, is the best that can be done at the present level of the present evolving theory. The derivations certainly demystify the concepts of intrinsic spacetime and intrinsic mass

introduced as *ansatz* in section 4 of [1]. There are, however, more formal and more complete derivations of these concepts along with the concepts of absolute intrinsic spacetime containing absolute intrinsic rest mass, which underlies absolute spacetime containing absolute rest mass and relativistic intrinsic spacetime containing relativistic intrinsic mass, which underlies relativistic spacetime containing relativistic mass, to be presented elsewhere with further development.

2 Validating perfect symmetry of state among the four universes isolated

Perfect symmetry of natural laws among the four universes namely, the positive universe, the negative universe, the positive time-universe and the negative time-universe, whose metric spacetimes and underlying intrinsic metric spacetimes are depicted in Figs. 8a and 8b of the first part of this paper [3], has been demonstrated in section 2 of that paper. Perfect symmetry of state among the universes shall now be demonstrated in this section. Perfect symmetry of state exists among the four universes if the masses of the four members of every quartet of symmetry-partner particles or objects in the four universes have identical magnitudes, shapes and sizes and if they perform identical relative motions in their universes at all times. These conditions shall be shown to be met in this section.

2.1 Identical magnitudes of masses and of shapes and sizes of the members of every quartet of symmetry-partner particles or objects in the four universes

As illustrated in Fig. 2a or 3a, the one-dimensional intrinsic rest mass (or proper nomass) ϕm_0 "projected" into the projective isotropic one-dimensional proper (or classical) intrinsic space (or proper nospace) $\phi\rho'$, lies directly underneath the three-dimensional rest mass m_0 in the proper (or classical) Euclidean 3-space Σ' of the positive (or our) universe with respect to 3-observers in Σ'. Likewise the "projective" one-dimensional intrinsic rest mass ϕm_0^0 in the projective one-dimensional isotropic proper intrinsic space $\phi\rho^{0\prime}$ lies directly underneath the three-dimensional rest mass m_0^0 in the proper (or classical) Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$ in Fig. 2b or 3b.

Now the rest mass m_0 is the outward (or physical) manifestation in the proper (or classical) physical Euclidean 3-space Σ' of the one-dimensional intrinsic rest mass ϕm_0 in the one-dimensional proper (or classical) intrinsic space $\phi\rho'$ lying underneath m_0 in Σ' in Fig. 2a or 3a. It then follows that m_0 and ϕm_0 are equal in magnitude, that is, $m_0 = |\phi m_0|$.

But the one-dimensional intrinsic rest mass ϕm_0 in $\phi\rho'$ along the horizontal is equal in magnitude to the one-dimensional rest mass m_0^0 in the one-dimensional space $\rho^{0\prime}$ along the vertical that 'projects' ϕm_0 contained in $\phi\rho'$ along the horizontal in Fig. 2a or 3a. That is, $m_0^0 = |\phi m_0|$. By combining this with $m_0 = |\phi m_0|$ derived in the preceding paragraph, we have the equality in magnitude of the three-dimensional rest mass m_0 of a particle or object in our proper Euclidean 3-space Σ' and the one-dimensional rest mass m_0^0 of the symmetry-partner particle or object in the one-dimensional proper (or classical) space $\rho^{0\prime}$ (with respect to 3-observers in Σ') in Fig. 2a or 3a. That is, $m_0 = m_0^0$.

Finally the one-dimensional rest mass m_0^0 of a particle or object in the one-dimensional proper space $\rho^{0\prime}$ along the vertical with respect to 3-observers in our proper Euclidean 3-space Σ' in Fig. 2a or 3a, is what 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe observe as three-dimensional rest mass m_0^0 of the particle or object in $\Sigma^{0\prime}$. Consequently the one-dimensional rest mass m_0^0 of the particle or object in $\rho^{0\prime}$ in Fig. 2a or 3a is equal in magnitude to the three-dimensional rest mass m_0^0 of the particle or object in the proper Euclidean 3-space $\Sigma^{0\prime}$. This is certainly so since the geometrical contraction of the Euclidean 3-space $\Sigma^{0\prime}$ to one-dimensional space $\rho^{0\prime}$ and the consequent geometrical contraction of the three-dimensional rest mass m_0^0 in $\Sigma^{0\prime}$ to one-dimensional rest mass m_0^0 in $\rho^{0\prime}$ with respect to 3-observers in our Euclidean 3-space Σ', does not alter the magnitude of the rest mass m_0^0.

In summary, we have derived the simultaneous relations $m_0 = |\phi m_0|$ and $m_0^0 = |\phi m_0|$, from which we have, $m_0 = m_0^0$ in the above. Also since m_0^0 in $\Sigma^{0\prime}$ is the outward manifestation of ϕm_0^0 in $\phi\rho^{0\prime}$ in Fig. 2b or 3b, we have the equality in magnitude of m_0^0 and ϕm_0^0, that is, $m_0^0 = |\phi m_0^0|$, which, along with $m_0^0 = |\phi m_0|$ derived above, gives $\phi m_0^0 = \phi m_0$. The conclusion then is that the rest mass m_0 of a particle or object in the proper Euclidean 3-space Σ' of our (or positive) universe with respect to 3-observers in Σ', is equal in magnitude to the rest mass m_0^0 of the symmetry-partner particle or object in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$. The one-dimensional intrinsic rest mass ϕm_0 of the particle or object in our proper intrinsic space $\phi\rho'$ underlying m_0 in Σ' in Fig. 2a, 3a or 4a is equal in magnitude to the intrinsic rest mass ϕm_0^0 of the symmetry-partner particle or object in the proper intrinsic space $\phi\rho^{0\prime}$ underlying m_0^0 in $\Sigma^{0\prime}$ in Fig. 2b, 3b or 4b.

By repeating the derivations done between the positive (or our) universe and the positive time-universe, which lead to the conclusion reached in the foregoing paragraph, between the negative universe and the negative time-universe, (which shall not be done here in order to conserve space), we are also led to the conclusion that the rest mass $-m_0^*$ of a particle or object in the proper Euclidean 3-space $-\Sigma'^*$ of the negative universe with respect to 3-observers in $-\Sigma'^*$, is equal in magnitude to the rest mass $-m_0^{0*}$ of the symmetry-partner particle or object in the proper Euclidean 3-space $-\Sigma^{0\prime*}$ of the negative time-universe with respect to 3-observers in $-\Sigma^{0\prime*}$. The one-dimensional intrinsic rest mass $-\phi m_0^*$ of the particle or object in the proper intrinsic space $-\phi\rho'^*$ of the negative universe underlying $-m_0^*$ in $-\Sigma'^*$, is equal in magnitude to the intrinsic rest mass $-\phi m_0^{0*}$ of the symmetry-partner particle or

object in the proper intrinsic space $-\phi\rho^{0\prime *}$ underlying $-m_0^{0*}$ in $-\Sigma^{0\prime *}$ in the negative time-universe.

The perfect symmetry of state between the positive (or our) universe and the negative universe prescribed in [1], remains a prescription so far. It implies that the rest mass m_0 of a particle or object in the proper Euclidean 3-space Σ' of the positive (or our) universe, is identical in magnitude to the rest mass $-m_0^*$ of the symmetry-partner particle or object in the proper Euclidean 3-space $-\Sigma'^*$ of the negative universe, that is, $m_0 = |-m_0^*|$. The corresponding (prescribed) perfect symmetry of state between positive time-universe and the negative time-universe likewise implies that the rest mass m_0^0 of a particle or object in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe is identical in magnitude to the rest mass$-m_0^{0*}$ of its symmetry-partner in the proper Euclidean 3-space $-\Sigma^{0\prime *}$ of the negative time-universe, that is, $m_0^0 = |-m_0^{0*}|$.

The equality of magnitudes of symmetry-partner rest masses, $m_0 = |-m_0^*|$, that follows from the prescribed perfect symmetry of state between the positive (or our) universe and the negative universe and $m_0^0 = |-m_0^{0*}|$ that follows from the prescribed symmetry of state between the positive time-universe and the negative time-universe, discussed in the foregoing paragraph, are possible of formal proof, as shall be presented elsewhere. By combining these with $m_0 = m_0^0$ and $-m_0^* = -m_0^{0*}$ derived from Figs. 2a and 2b above, we obtain the equality of magnitudes of the rest masses of the four symmetry-partner particles or objects in the four universes, that is, $m_0 = |-m_0^*| = m_0^0 = |-m_0^{0*}|$. Consequently there is equality of magnitudes of the intrinsic rest masses in the one-dimensional intrinsic spaces of the quartet of symmetry-partner particles or objects in the four universes, that is, $|\phi m_0| = |-\phi m_0^*| = |\phi m_0^0| = |-\phi m_0^{0*}|$.

Having demonstrated the equality of magnitudes of the rest masses of the members of every quartet of symmetry-partner particles or objects in the four universes, (to the extent that $m_0 = |-m_0^*|$ between the positive (or our) universe and the negative universe and $m_0^0 = |-m_0^{0*}|$ between the positive and negative time-universes are valid), let us also show their identical shapes and sizes.

Now the rest mass m_0 being the outward manifestation in our proper Euclidean 3-space Σ' of the intrinsic rest mass ϕm_0 of intrinsic length $\Delta\phi\rho'$ in the one-dimensional proper intrinsic space $\phi\rho'$ and the three-dimensional rest mass m_0^0 in the proper Euclidean 3-space $\Sigma^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$, being what geometrically contracts to the one-dimensional rest mass m_0^0 of length $\Delta\rho^{0\prime}$ in $\rho^{0\prime}$ with respect to 3-observers in our Euclidean 3-space Σ' and since $\Delta\rho^{0\prime}$ along the vertical projects $\Delta\phi\rho'$ into Σ' along the horizontal, then the length $\Delta\rho^{0\prime}$ of the one-dimensional rest mass m_0^0 in $\rho^{0\prime}$ has the same magnitude as the intrinsic length $\Delta\phi\rho'$ of the intrinsic rest mass ϕm_0 in $\phi\rho'$, that is, $\Delta\rho^{0\prime} = |\Delta\phi\rho'|$. Consequently the volume $\Delta\Sigma^{0\prime}$ of the Euclidean 3-space $\Sigma^{0\prime}$ occupied by the three-dimensional rest mass m_0^0 with respect to 3-observers in $\Sigma^{0\prime}$ has the same magnitude as the volume $\Delta\Sigma'$ of the Euclidean 3-space Σ' occupied by the rest mass m_0 with respect to 3-observers in Σ'; $\Delta\Sigma'$ occupied by m_0 being the outward manifestation of $\Delta\phi\rho'$ occupied by ϕm_0. In other words, the rest mass m_0 in Σ' has the same size as its symmetry-partner m_0^0 in $\Sigma^{0\prime}$.

Further more, the shape of the outward manifestation of ϕm_0 in the proper Euclidean 3-space Σ', that is, the shape of m_0 in Σ', with respect to 3-observers in Σ', is the same as the shape of the three-dimensional rest mass m_0^0 in the proper Euclidean 3-space $\Sigma^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$. In providing justification for this, let us recall the discussion leading to Fig. 6a and 6b of [1], that the intrinsic rest masses ϕm_0 of particles and objects, which appear as lines of intrinsic rest masses along the one-dimensional isotropic proper intrinsic space $\phi\rho'$ relative to 3-observers in the proper Euclidean 3-space Σ', as illustrated for a few objects in Fig. 6a of [1], are actually three-dimensional intrinsic rest masses ϕm_0 in three-dimensional proper intrinsic space $\phi\Sigma'$ with respect to three-dimensional intrinsic-rest-mass-observers (or 3-intrinsic-observers) in $\phi\Sigma'$, as also illustrated for a few objects in Fig. 6b of [1]. The shape of the three-dimensional intrinsic rest mass ϕm_0 of an object or particle in the three-dimensional intrinsic space $\phi\Sigma'$ with respect to 3-intrinsic-observers in $\phi\Sigma'$, is the same as the shape of its outward manifestation in the proper Euclidean 3-space Σ', that is, the same as the shape of the rest mass m_0 in Σ', with respect to 3-observers in Σ'.

Since the line of intrinsic rest mass ϕm_0 in one-dimensional proper intrinsic space $\phi\rho'$ relative to 3-observers in Σ', (which is a three-dimensional intrinsic rest mass ϕm_0 in three-dimensional proper intrinsic space $\phi\Sigma'$ with respect to 3-intrinsic-observers in $\phi\Sigma'$), is the projection along the horizontal of the line of rest mass m_0^0 in the one-dimensional proper space $\rho^{0\prime}$ along the vertical relative to 3-observers in our proper Euclidean 3-space Σ', (which is a 3-dimensional rest mass m_0^0 in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in $\Sigma^{0\prime}$), the shape of the three-dimensional intrinsic rest mass ϕm_0 in $\phi\Sigma'$ with respect to 3-intrinsic-observers in $\phi\Sigma'$ is the same as the shape of the three-dimensional rest mass m_0^0 in $\Sigma^{0\prime}$ with respect to 3-observers in $\Sigma^{0\prime}$. It then follows from this and the conclusion (that the shape of ϕm_0 in $\phi\Sigma'$ is the same as the shape of m_0 in Σ') reached in the preceding two paragraphs, that the shapes of the rest masses m_0 in our proper Euclidean 3-space Σ' and m_0^0 in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe are the same, as stated earlier.

The identical sizes and shapes of the the rest mass m_0 of a particle or object in the proper Euclidean 3-space Σ' of our universe and of the rest mass m_0^0 of its symmetry-partner in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe, concluded from the foregoing, is equally true between the rest mass $-m_0^*$ in the Euclidean 3-space $-\Sigma'^*$ of the negative universe and its symmetry-partner $-m_0^{0*}$ in the

Euclidean 3-space $-\Sigma^{0\prime *}$ of the negative time-universe.

When the preceding paragraph is combined with the identical shapes and sizes of the rest mass m_0 of a particle or object in the proper Euclidean 3-space Σ' of the positive (or our) universe and of the rest mass $-m_0^*$ of its symmetry-partner in the proper Euclidean 3-space $-\Sigma'^*$ of the negative universe, which the so far prescribed perfect symmetry of state between our universe and the negative universe implies, as well as the identical shapes and sizes of the rest mass m_0^0 of a particle or object in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe and of the rest mass $-m_0^{0*}$ of its symmetry-partner in the proper Euclidean 3-space $-\Sigma^{0\prime *}$ of the negative time-universe, which the so far prescribed perfect symmetry of state between positive time-universe and the negative time-universe implies, we have the identical shapes and sizes of the four members of the quartet of symmetry-partner particles or objects in the four universes, and this is true for every such quartet of symmetry-partner particles or objects.

2.2 Perfect symmetry of relative motions always among the members of every quartet of symmetry-partner particles or objects in the four universes

As mentioned at the beginning of this section, the second condition that must be met for symmetry of state to obtain among the four universes isolated in part one of this paper [3] and illustrated in Figs. 8a and 8b of that paper namely, the positive (or our) universe, the negative universe, the positive time-universe and the negative time-universe, is that the members of every quartet of symmetry-partner particles or objects in the universes, now shown to have identical magnitudes of masses, identical sizes and identical shapes, are involved in identical motions relative to identical symmetry-partner observers or frames of reference in the universes at all times. The *reductio ad absurdum* method of proof shall be applied to show that this second condition is also met. We shall assume that the quartet of symmetry-partner particles or objects in the four universes are not involved in identical relative motions and show that this leads to a violation of Lorentz invariance.

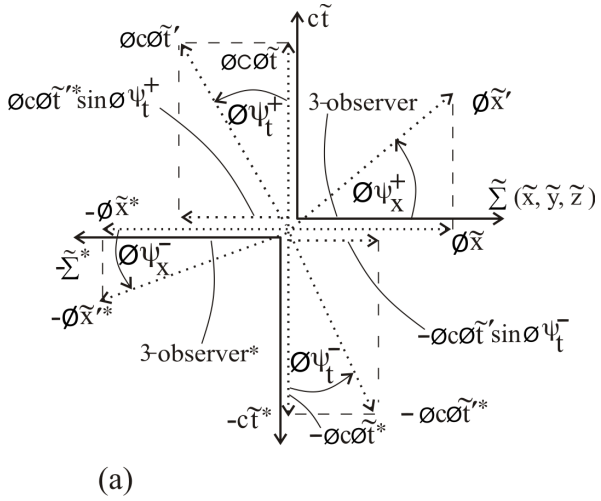
Let us start with the assumption that the members of a quartet of symmetry-partner particles or objects in the four universes are in arbitrary motions at different speeds relative to the symmetry-partner observers of frames of reference in their respective universes at every given moment. This assumption implies that given an object on earth in our universe in motion at a speed v_{x^+} along the north pole of the earth, say, relative to our earth at a given instant, then its symmetry-partner on earth in the negative universe is in motion at a speed v_{x^-} along the north pole relative to the earth of the negative universe at the same instant; the symmetry-partner object on earth in the positive time-universe is motion at a speed $v_{x^{0+}}$ along the north pole relative to the earth of the positive time-universe at the same instant and the symmetry-partner object on earth in the negative time-universe is in motion at a speed $v_{x^{0-}}$ along the north pole relative to the earth of the negative time-universe at the same instant, where it is being assumed that the speeds v_{x^+}, v_{x^-}, $v_{x^{0+}}$ and $v_{x^{0-}}$ have different magnitudes and each could take on arbitrary values lower than c, including zero. They may as well be assumed to be moving along arbitrary directions on earths in their respective universes.

The geometrical implication of the assumption made in the foregoing paragraph is that the equal intrinsic angle $\phi\psi$ of relative rotations of intrinsic affine space and intrinsic affine time coordinates in the four quadrants, drawn upon the proper (or classical) metric spacetimes/intrinsic spacetimes of the positive (or our) universe and the negative universe in Fig. 8a of [3], as Fig. 10a of that paper, and upon the proper (or classical) metric spacetimes/intrinsic spacetimes of the positive time-universe and negative time-universe in Fig. 8b of [3], as Fig. 10b of that paper, will take on different values $\phi\psi_x^+$, $\phi\psi_x^-$, $\phi\psi_t^+$, $\phi\psi_t^-$, $\phi\psi_{x^0}^+$, $\phi\psi_{x^0}^-$, $\phi\psi_{t^0}^+$ and $\phi\psi_{t^0}^-$ as depicted in Figs. 5a and 5b.

The rotations of $\phi\tilde{x}'$ by intrinsic angle $\phi\psi_x^+$ relative to $\phi\tilde{x}$ along the horizontal in the first quadrant and the rotation of $\phi c\phi\tilde{t}'$ by intrinsic angle $\phi\psi_t^+$ relative to $\phi c\phi\tilde{t}$ along the vertical in the second quadrant are valid with respect to the 3-observer in $\tilde{\Sigma}$ in Fig. 5a, where $\sin\phi\psi_x^+ = \phi v_{x^+}/\phi c$ and $\sin\phi\psi_t^+ = \phi v_{t^+}/\phi c$. On the other hand, the rotation of $-\phi\tilde{x}'^*$ at intrinsic angle $\phi\psi_x^-$ relative to $-\phi\tilde{x}^*$ along the horizontal in the third quadrant and the rotation of $-\phi c\phi\tilde{t}'^*$ by intrinsic angle $\phi\psi_t^-$ relative to $-\phi c\phi\tilde{t}^*$ along the vertical in the fourth quadrant in Fig. 5a are valid with respect to the 3-observer* in $-\tilde{\Sigma}^*$, where $\sin\phi\psi_x^- = \phi v_{x^-}/\phi c$ and $\sin\phi\psi_t^- = \phi v_{t^-}/\phi c$.

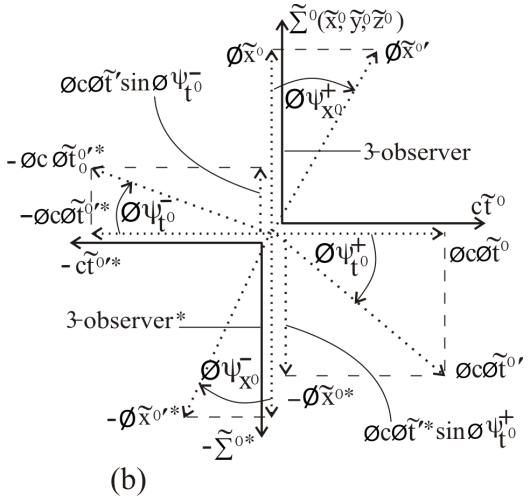
The rotations of $\phi\tilde{x}^{0\prime}$ by intrinsic angle $\phi\psi_{x^{0+}}$ relative to $\phi\tilde{x}^0$ along the vertical in the first quadrant and the rotation of $\phi c\phi\tilde{t}^{0\prime}$ by intrinsic angle $\phi\psi_{t^{0+}}$ relative to $\phi c\phi\tilde{t}^0$ along the horizontal in the fourth quadrant are valid with respect to the 3-observer in $\tilde{\Sigma}^0$ in Fig. 5b, where $\sin\phi\psi_{x^0}^+ = \phi v_{x^{0+}}/\phi c$ and $\sin\phi\psi_{t^0}^+ = \phi v_{t^{0+}}/\phi c$. On the other hand, the rotation of $-\phi\tilde{x}^{0\prime *}$ at intrinsic angle $\phi\psi_{x^0}^-$ relative to $-\phi\tilde{x}^{0*}$ along the vertical in the third quadrant and the rotation of $-\phi c\phi\tilde{t}^{0\prime *}$ by intrinsic angle $\phi\psi_{t^0}^-$ relative to $-\phi c\phi\tilde{t}^{0*}$ along the vertical in the second quadrant in Fig. 5b are valid with respect to the 3-observer* in $-\tilde{\Sigma}^{0*}$, where $\sin\phi\psi_{x^0}^- = \phi v_{x^{0-}}/\phi c$ and $\sin\phi\psi_{t^0}^- = \phi v_{t^{0-}}/\phi c$.

Although the intrinsic angles $\phi\psi_x^+$, $\phi\psi_x^-$, $\phi\psi_t^+$ and $\phi\psi_t^-$, which are related to the intrinsic speeds ϕv_{x^+}, ϕv_{x^-}, ϕv_{t^+} and ϕv_{t^-}, as $\sin\phi\psi_x^+ = \phi v_{x^+}/\phi c$; $\sin\phi\psi_x^- = \phi v_{x^-}/\phi c$; $\sin\phi\psi_t^+ = \phi v_{t^+}/\phi c$; and $\sin\phi\psi_t^- = \phi v_{t^-}/\phi c$ respectively in Fig. 5a, are different in magnitude as being assumed and although the intrinsic angles $\phi\psi_{x^0}^+$, $\phi\psi_{x^0}^-$, $\phi\psi_{t^0}^+$, $\phi\psi_{t^0}^-$, which are related to intrinsic speeds $\phi v_{x^{0+}}$, $\phi v_{x^{0-}}$, $\phi v_{t^{0+}}$ and $\phi v_{t^{0-}}$ as $\sin\phi\psi_{x^0}^+ = \phi v_{x^{0+}}/\phi c$; $\sin\phi\psi_{x^0}^- = \phi v_{x^{0-}}/\phi c$; $\sin\phi\psi_{t^0}^+ = \phi v_{t^{0+}}/\phi c$; and $\sin\phi\psi_{t^0}^- = \phi v_{t^{0-}}/\phi c$ respectively in Fig. 5b, are different in magnitude as being assumed, it must be remembered that the intrinsic angles $\phi\psi_x^+$, $\phi\psi_x^-$, $\phi\psi_{t^+}$, $\phi\psi_t^-$ in Fig. 5a are equal to the intrin-



(a)

Fig. 5: (a) Rotations of intrinsic affine spacetime coordinates of intrinsic particle's frame relative to intrinsic observer's frame due to assumed non-symmetrical motions of symmetry-partner particles relative to symmetry-partner observers in the four universes, with respect to 3-observers in the Euclidean 3-spaces in the positive (or our) universe and the negative universe.



(b)

Fig. 5: (b) Rotations of intrinsic affine spacetime coordinates of intrinsic particle's frame relative to intrinsic observer's frame due to assumed non-symmetrical motions of symmetry-partner particles relative to symmetry-partner observers in the four universes, with respect to 3-observers in the Euclidean 3-spaces in the positive time-universe and the negative time-universe.

sic angles $\phi\psi^+_{t^0}$, $\phi\psi^-_{t^0}$, $\phi\psi^+_{x^0}$, $\phi\psi^-_{x^0}$ respectively in Fig. 5b.

That is, $\phi\psi^+_x = \phi\psi^+_{t^0}$; $\phi\psi^-_x = \phi\psi^-_{t^0}$; $\phi\psi^+_t = \phi\psi^+_{x^0}$ and $\phi\psi^-_t = \phi\psi^-_{x^0}$ in Figs. 5a and 5b. Consequently the intrinsic speeds ϕv_{x^+}, ϕv_{x^-}, ϕv_{t^+} and ϕv_{t^-} in Fig. 5a are equal to $\phi v_{t^{0+}}$, $\phi v_{t^{0-}}$, $\phi v_{x^{0+}}$ and $\phi v_{x^{0-}}$ respectively in Fig. 5b. That is, $\phi v_{x^+} = \phi v_{t^{0+}}$; $\phi v_{x^-} = \phi v_{t^{0-}}$; $\phi v_{t^+} = \phi v_{x^{0+}}$ and $\phi v_{t^-} = \phi v_{x^{0+}}$ in Figs. 5a and 5b.

By following the procedure used to derive partial intrinsic Lorentz transformation with respect to the 3-observer in $\tilde{\Sigma}$ from Fig. 8a of [1], the unprimed intrinsic affine coordinate $\phi\tilde{x}$ along the horizontal is the projection of the inclined $\phi\tilde{x}'$ in the first quadrant in Fig. 5a. That is, $\phi\tilde{x} = \phi\tilde{x}' \cos\phi\psi^+_x$. Hence we can write

$$\phi\tilde{x}' = \phi\tilde{x} \sec\phi\psi^+_x$$

This is all the intrinsic coordinate transformation that could have been possible with respect to the 3-observer in $\tilde{\Sigma}$ along the horizontal in the first quadrant in Fig. 5a, but for the fact that the inclined negative intrinsic coordinate $-\phi c\phi\tilde{t}'^*$ of the negative universe in the fourth quadrant also projects a component $-\phi c\phi\tilde{t}' \sin\phi\psi^-_t$ along the horizontal, which must be added to the right-hand side of the last displayed equation to have

$$\phi\tilde{x}' = \phi\tilde{x} \sec\phi\psi^+_x - \phi c\phi\tilde{t}' \sin\phi\psi^-_t; \qquad (*)$$

w.r.t 3 − observer in $\tilde{\Sigma}$.

As mentioned in the derivation of $(*)$, but for $\phi\psi^+_x = \phi\psi^-_x = \phi\psi$ with Fig. 8a in [1], the dummy star label on the component $-\phi c\phi\tilde{t}'^* \sin\phi\psi^-_t$ projected along the horizontal has been removed, since the projected component is now an intrinsic coordinate in the positive universe.

But $\phi c\phi\tilde{t} = \phi c\phi\tilde{t}' \cos\phi\psi^+_t$ or $\phi c\phi\tilde{t}' = \phi c\phi\tilde{t} \sec\phi\psi^+_t$ along the vertical in the second quadrant in the same Fig. 5a. By replacing $\phi c\phi\tilde{t}'$ by $\phi c\phi\tilde{t} \sec\phi\psi^+_t$ at the right-hand side of $(*)$ we have

$$\phi\tilde{x}' = \phi\tilde{x} \sec\phi\psi^+_x - \phi c\phi\tilde{t} \sec\phi\psi^+_t \sin\phi\psi^-_t; \qquad (9)$$

w.r.t 3 − observer in $\tilde{\Sigma}$. Eq. (9) is the final form of the partial intrinsic Lorentz transformation that the 3-observer in $\tilde{\Sigma}$ in our universe could derive along the horizontal in the first quadrant from Fig. 5a.

By applying the same procedure used to derive Eq. (9) from the first and fourth quadrants of Fig. 5a to the first and second quadrants of Fig. 5b, the counterpart of Eq. (9) that is valid with respect to the 3-observer in $\tilde{\Sigma}^0$ in that figure is the following:

$$\phi\tilde{x}^{0\prime} = \phi\tilde{x}^0 \sec\phi\psi^+_{x^0} - \phi c\phi\tilde{t}^0 \sec\phi\psi^+_{t^0} \sin\phi\psi^-_{t^0}; \qquad (10)$$

w.r.t 3 − observer in $\tilde{\Sigma}^0$. Again Eq. (10) is the final form of the partial intrinsic Lorentz transformation that the 3-observer in $\tilde{\Sigma}^0$ in the positive time-universe could derive along the vertical in the first quadrant from Fig. 5b. By collecting Eqs. (9) and (10) we have

$$\left.\begin{array}{rcl} \phi\tilde{x}' & = & \phi\tilde{x} \sec\phi\psi^+_x - \phi c\phi\tilde{t} \sec\phi\psi^+_t \sin\phi\psi^-_t; \\ & & (\text{w.r.t } 3 - \text{observer in } \tilde{\Sigma}); \\ \phi\tilde{x}^{0\prime} & = & \phi\tilde{x}^0 \sec\phi\psi^+_{x^0} - \phi c\phi\tilde{t}^0 \sec\phi\psi^+_{t^0} \sin\phi\psi^-_{t^0}; \\ & & (\text{w.r.t } 3 - \text{observer in } \tilde{\Sigma}^0) \end{array}\right\}. \qquad (11)$$

However system (11) is useless because it is neither the full intrinsic Lorentz transformation in our (or positive) universe nor in the the positive time-universe. This is so because the second equation of system (11) contains intrinsic coordinates of the positive time-universe, which are elusive to observers in our universe or which cannot appear in physics in our universe. On the other hand, the first equation contains the intrinsic spacetime coordinates of our universe, which cannot appear in physics in the positive time-universe.

In order to make system (11) a valid full intrinsic spacetime coordinate transformation (i.e. to make it full intrinsic Lorentz transformation) in our universe, we must transform the intrinsic spacetime coordinates of the positive time-universe in the second equation into the intrinsic spacetime coordinates of our universe. As derived in part one of this paper [3], we must let $\phi\tilde{x}^{0\prime} \to \phi c\phi\tilde{t}'$, $\phi\tilde{x}^0 \to \phi c\phi\tilde{t}$ and $\phi c\phi\tilde{t}^0 \to \phi\tilde{x}$ in the second equation of system (11), thereby converting system (11) to the following

$$\left.\begin{array}{rcl} \phi\tilde{x}' & = & \phi\tilde{x} \sec\phi\psi^+_x - \phi c\phi\tilde{t} \sec\phi\psi^+_t \sin\phi\psi^-_t; \\ & & (\text{w.r.t } 3 - \text{observer in } \tilde{\Sigma}); \\ \phi c\phi\tilde{t}' & = & \phi c\phi\tilde{t} \sec\phi\psi^+_t - \sec\phi\psi^+_x \sin\phi\psi^-_x; \\ & & (\text{w.r.t } 1 - \text{observer in } c\tilde{t}) \end{array}\right\}. \qquad (12)$$

Fig. 5b cannot serve the role of a complementary diagram to Fig. 5a because it contains spacetime and intrinsic spacetime coordinates of the positive time-universe and negative time-universe that are elusive to observers in our universe and negative universe. This has been discussed for Figs. 10a and 10b of [3]. In order to make Fig. 5b a valid complementary diagram to Fig. 5a, the spacetime/intrinsic spacetime coordinates of the positive and negative time-universes in it must be transformed into those of our universe and the negative universe, as done between Fig. 10b and Fig. 11a of [3], to have Fig. 5c.

Fig. 5c containing spacetime/intrinsic spacetime coordinates of the positive (or our) universe and negative universe (obtained from Fig. 5b) is now a valid complementary diagram to Fig. 5a for the purpose of deriving the ϕLT/LT in our universe and negative universe. Observe that the 3-observers in the Euclidean 3-spaces $\tilde{\Sigma}^0$ and $-\tilde{\Sigma}^{0*}$ of the positive and negative time-universes in Fig. 5b have transformed into 1-observers in the time dimensions $c\tilde{t}$ and $-c\tilde{t}^*$ of our universe and the negative universe in Fig. 5c.

The second equation of system (12) has been derived from Fig. 5c with respect to the 1-observer in the time dimension $c\tilde{t}$

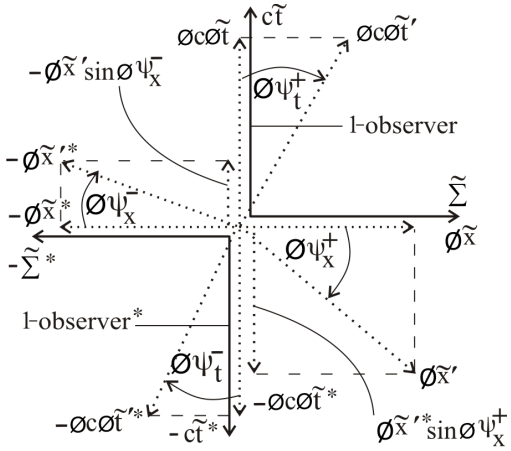


Fig. 5: (c) Complementary diagram to Fig. 5a obtained by transforming the spacetime/intrinsic spacetime coordinates of the positive time-universe and the negative time-universe in Fig. 5b into the spacetime/intrinsic spacetime coordinates of our universe and the negative universe.

in that diagram. It is a valid complementary partial intrinsic spacetime transformation to the first equation of system (12) or to Eq. (10) derived with respect to 3-observer in the Euclidean 3-space $\tilde{\Sigma}$ from Fig. 5a. Thus system (12) is the complete intrinsic Lorentz transformation derivable from Figs. 5a and 5c with respect to 3-observer in $\tilde{\Sigma}$ and 1-observer in $c\tilde{t}$.

By using the definitions given earlier namely,

$$\begin{aligned}
\sin\phi\psi_x^+ &= \sin\phi\psi_{t^0}^+ = \phi v_{x^+}/\phi c;\\
\sin\phi\psi_x^- &= \sin\phi\psi_{t^0}^- = \phi v_{x^-}/\phi c;\\
\sin\phi\psi_{x^0}^+ &= \sin\phi\psi_t^+ = \phi v_{t^+}/\phi c \ \text{ and}\\
\sin\phi\psi_{x^0}^- &= \sin\phi\psi_t^- = \phi v_{t^-}/\phi c;
\end{aligned}$$

system (12) is given explicitly in terms of intrinsic speeds as follows:

$$\left.\begin{aligned}
\phi\tilde{x}' &= \left(1-\frac{\phi v_{x^+}^2}{\phi c^2}\right)^{-\frac{1}{2}}\phi\tilde{x}-\\
&\quad -\left(1-\frac{\phi v_{t^+}^2}{\phi c^2}\right)^{-\frac{1}{2}}(\phi v_{t^-})\phi\tilde{t};\\
&\quad (\text{w.r.t. } 3-\text{observer in } \tilde{\Sigma})\\
\phi\tilde{t}' &= \left(1-\frac{\phi v_{t^+}^2}{\phi c^2}\right)^{-\frac{1}{2}}\phi\tilde{t}-\\
&\quad -\left(1-\frac{\phi v_{x^+}^2}{\phi c^2}\right)^{-\frac{1}{2}}\frac{\phi v_{x^-}}{\phi c^2}\phi\tilde{x};\\
&\quad (\text{w.r.t. } 1-\text{observer in } c\tilde{t})
\end{aligned}\right\} \qquad (13)$$

The outward manifestation on the flat four-dimensional spacetime of systems (12) and (13) are given respectively as follows:

$$\left.\begin{aligned}
\tilde{x}' &= \tilde{x}\sec\psi_x^+ - c\tilde{t}\sec\psi_t^+\sin\psi_t^-;\\
&\quad \tilde{y}'=\tilde{y};\ \tilde{z}'=\tilde{z};\\
&\quad (\text{w.r.t } 3-\text{observer in } \tilde{\Sigma});\\
c\tilde{t}' &= c\tilde{t}\sec\psi_t^+ - \tilde{x}\sec\psi_x^+\sin\psi_x^-;\\
&\quad (\text{w.r.t } 1-\text{observer in } c\tilde{t})
\end{aligned}\right\} \qquad (14)$$

and

$$\left.\begin{aligned}
\tilde{x}' &= \left(1-\frac{v_{x^+}^2}{c^2}\right)^{-\frac{1}{2}}\tilde{x}-\left(1-\frac{v_{t^+}^2}{c^2}\right)^{-\frac{1}{2}}(v_{t^-})\tilde{t};\\
&\quad \tilde{y}'=\tilde{y};\ \tilde{z}'=\tilde{z};\\
&\quad (\text{w.r.t. } 3-\text{observer in } \tilde{\Sigma})\\
\tilde{t}' &= \left(1-\frac{v_{t^+}^2}{c^2}\right)^{-\frac{1}{2}}\tilde{t}-\left(1-\frac{v_{x^+}^2}{c^2}\right)^{-\frac{1}{2}}\frac{v_{x^-}}{c^2}\tilde{x};\\
&\quad (\text{w.r.t. } 1-\text{observer in } c\tilde{t})
\end{aligned}\right\}. \qquad (15)$$

As can be easily shown, system (12) or (13) contradicts (or does not lead to) intrinsic Lorentz invariance (ϕLI) for $\phi\psi_x^+ \neq \phi\psi_x^- \neq \phi\psi_t^+ \neq \phi\psi_t^-$ (or for $\phi v_{x^+} \neq \phi v_{x^-} \neq \phi v_{t^+} \neq \phi v_{t^-}$). System (14) or (15) likewise does not lead to Lorentz invariance (LI) for $\psi_x^+ \neq \psi_x^- \neq \psi_t^+ \neq \psi_t^-$ (or $v_{x^+} \neq v_{x^-} \neq v_{t^+} \neq v_{t^-}$). Even if only one of the four intrinsic angles $\phi\psi_x^+$, $\phi\psi_x^-$, $\phi\psi_t^+$ and $\phi\psi_t^-$ is different from the rest (or if only one of the four intrinsic speeds ϕv_{x^+}, ϕv_{x^-}, ϕv_{t^+} and ϕv_{t^-} is different from the rest), system (12) or (13) still contradicts ϕLI. And even if only one of the four angles ψ_x^+, ψ_x^-, ψ_t^+ and ψ_t^- is different from the rest (or if only one of the four speeds v_{x^+}, v_{x^-}, v_{t^+} and v_{t^-} is different from the rest), system (14) or (15) still contradicts the LI.

The assumption made initially that members of a quartet of symmetry-partner particles or objects in the four universes

are in non-symmetrical relative motions in their universes, which gives rise to Fig. 5a-c, has led to the non-validity of intrinsic Lorentz invariance in intrinsic special relativity (ϕSR) and of Lorentz invariance in special relativity (SR) in our universe and indeed in the four universes. This invalidates the initial assumption, since Lorentz invariance is immutable on the flat four-dimensional spacetime of the special theory of relativity. The conclusion then is that all the four members of every quartet of symmetry-partner particles or objects in the four universes are in identical (or symmetrical) relative motions at all times.

Having shown that the members of every quartet of symmetry-partner particles or objects in the four universes have identical magnitudes of masses, identical shapes and identical sizes, (in so far as the prescribed identical magnitudes of masses, identical shapes and identical sizes of symmetry-partner particles or objets in the positive (or our) universe and the negative universe is valid), in the preceding sub-section and that they are involved in identical relative motions at all times in this sub-section, the perfect symmetry of state among the four universes has been demonstrated. Although gravity is being assumed to be absent in this and the previous papers [1-3], it is interesting to note that gravitational field sources of identical magnitudes of masses, identical sizes and identical shapes, which hence give rise to identical gravitational fields, are located at symmetry-partner positions in spacetimes in the four universes.

3 Summary and conclusion

This section is for the two parts of the initial paper [1] and [2], this paper and its first part [3]. The co-existence in nature of four symmetrical universes identified as positive (or our) universe, negative universe, positive time-universe and negative time-universe in different spacetime/intrinsic spacetime domains, have been exposed in these papers. The four universes exhibit perfect symmetry of natural laws and perfect symmetry of state. This implies that natural laws take on identical forms in the four universes and that all members of every quartet of symmetry-partner particles or objects in the four universes have identical magnitudes of masses, identical shapes and identical sizes and that they are involved in identical relative motions in their universes at all times, as demonstrated. The four universes constitute a four-world background to the special theory of relativity in each universe.

The flat two-dimensional intrinsic spacetime of the intrinsic special theory of relativity (ϕSR), containing one-dimensional intrinsic masses of particles and objects in one-dimensional intrinsic space, which underlies the flat four-dimensional spacetime of the special theory of relativity (SR) containing three-dimensional masses of particles and objects in Euclidean 3-space in each universe, introduced (as *ansatz* in the first paper [1], is isolated in this fourth paper. The two-dimensional intrinsic spacetime is indispensable in special relativity/intrinsic special relativity (SR/ϕSR) in the four-world picture, because the new set of spacetime/intrinsic spacetime diagrams for deriving Lorentz transformation/intrinsic Lorentz transformation (LT/ϕLT) and their inverses in the four-world picture, involve relative rotations of intrinsic spacetime coordinates of two frames in relative motion.

The LT/ϕLT and their inverses are derived from a new set of spacetime/intrinsic spacetime diagrams on the combined spacetimes/intrinsic spacetimes of the positive (or our) universe and the negative universe as one pair of universes and on combined spacetimes/intrinsic spacetimes of the positive time-universe and the negative time-universe as another pair of universes. The two pairs of spacetimes/intrinsic spacetimes co-exist in nature, consequently the spacetime/intrinsic spacetime diagram drawn on one pair co-exists with and must complement the spacetime/intrinsic spacetime diagram drawn on the other pair in deriving the LT/ϕLT and their inverses (with a set of four diagrams in all) in each universe, as done in the first paper [1] and validated formally in the third paper [3].

The proper (or classical) Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe with respect to 3-observers in it, is what appears as the proper time dimension ct' of the positive (or our) universe relative to 3-observers in the proper Euclidean 3-space Σ' of our universe and the proper Euclidean 3-space Σ' of the positive (or our) universe with respect to 3-observers in it, is what appears as the proper time dimension $ct^{0\prime}$ of the positive time-universe relative to 3-observers in the proper Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe. The proper Euclidean 3-space $-\Sigma^{0\prime *}$ of the negative time-universe is likewise the proper time dimension $-ct'^{*}$ of the negative universe and the proper Euclidean 3-space $-\Sigma'^{*}$ of the negative universe is the proper time dimension $-ct^{0\prime *}$ of the negative time-universe. The important revelation in this is that time is not a fundamental (or "created") concept, but a secondary concept that evolved from the concept of space. Time dimension does not exist in an absolute sense, as does 3-space, but in a relative sense.

The positive time-universe cannot be perceived better than the time dimension ct' of the positive (or our) universe by 3-observers in the Euclidean 3-space Σ' of our universe and the negative time-universe cannot be perceived better than the time dimension $-ct'^{*}$ of the negative universe by 3-observers in the Euclidean 3-space $-\Sigma'^{*}$ of the negative universe. Conversely, the positive (or our) universe cannot be perceived better than the time dimension $ct^{0\prime}$ of the positive time-universe by 3-observers in the Euclidean 3-space $\Sigma^{0\prime}$ of the positive time-universe and the negative universe cannot be perceived better than the time dimension $-ct^{0\prime *}$ of the negative time-universe by 3-observers in the Euclidean 3-space $-\Sigma^{0\prime *}$ of the negative time-universe. It can thus be said that the positive time-universe and the negative time-universe are imperceptibly hidden in the time dimensions of the positive (or our) universe and the negative universe respectively rela-

tive to 3-observers in the Euclidean 3-spaces in our universe and the negative universe and conversely.

Physicists in our (or positive) universe and negative universe can formulate special relativity and special-relativistic physics in general in terms of the spacetime/intrinsic spacetime dimensions (or coordinates) and physical parameters/intrinsic parameters of our universe and the negative universe only. Physicists in the positive time-universe and the negative time-universe can likewise formulate special relativity and special-relativistic physics in general in terms of the spacetime/intrinsic spacetime dimensions (or coordinates) and parameters/intrinsic parameters of the positive and the negative time-universes only. It is to this extent that it can still be said that special relativity and special-relativistic physics in general, pertain to a two-world background, knowing that the two-world picture actually encompasses four universes; two of them being imperceptibly hidden in the time dimensions.

Experimental validation ultimately of the co-existence in nature of four symmetrical universes will give a second testimony to their isolation theoretically in these papers. The next natural step is to investigate the possibility of subsuming the theory of relativistic gravity into the four-world picture.

Acknowledgements

Grateful acknowledgement is made to Dr. Ibukun Omotehinse and Mr. Cyril Ugwu for fully bearing the costs of publication of this and the the previous three papers.

Submitted on February 27, 2010 / Accepted on March 01, 2010

References

1. Adekugbe A. O. J. Two-world background of special relativity. Part I. *Progress in Physics*, 2010, v. 1, 30–48.
2. Adekugbe A. O. J. Two-world background of special relativity. Part II. *Progress in Physics*, 2010, v. 1, 49–61.
3. Adekugbe A. O. J. Re-identification of the many-world background of special relativity as four-world background. Part I. *Progress in Physics*, 2011, v. 1, 3–24.

Missing Measurements of Weak-Field Gravity

Richard J. Benish

Eugene, Oregon, USA. E-mail: rjbenish@teleport.com

For practical and historical reasons, most of what we know about gravity is based on observations made or experiments conducted *beyond the surfaces* of dominant massive bodies. Although the *force* of gravity inside a massive body can sometimes be measured, it remains to demonstrate the *motion* that would be caused by that force through the body's center. Since the idea of doing so has often been discussed as a thought experiment, we here look into the possibility of turning this into a real experiment. Feasibility is established by considering examples of similar experiments whose techniques could be utilized for the present one.

1 Introduction

A recent paper in this journal (M. Michelini [1]) concerned the absence of measurements of Newton's constant, G, within a particular range of vacuum pressures. Important as it may be to investigate the physical reasons for this, a gap of equal, if not greater importance concerns the absence of gravity experiments that probe the *motion* of test objects through the centers of larger massive bodies. As is the case for most measurements of G, the apparatus for the present experimental idea is also a variation of a torsion balance. Before describing the modifications needed so that a torsion balance can measure through-the-center motion, let's consider the context in which we find this gap in experimentation.

Often found in undergraduate physics texts [2–5] is the following problem, discussed in terms of Newtonian gravity: A test object is dropped into an evacuated hole spanning a diameter of an otherwise uniformly dense spherical mass. One of the reasons this problem is so common is that the answer, the predicted equation of motion of the test object, is yet another instance of *simple harmonic motion*. What is rarely pointed out, however, is that we presently lack direct empirical evidence to verify the theoretical prediction. Confidence in the prediction is primarily based on the success of Newton's theory for phenomena that test the *exterior* solution. Extrapolating Newton's law to the interior is a worthwhile mathematical excercise. But a theoretical extrapolation is of lesser value than an empirical fact.

Essentially the same prediction follows from general relativity [6–9]. In this context too, the impression is sometimes given that the predicted effect is a physical fact. A noteworthy example is found in John A. Wheeler's book, *"A Journey into Gravity and Spacetime"*, in which he refers to the phenomenon as "boomeranging". Wheeler devotes a whole (10-page) chapter to the subject because, as he writes, "Few examples of gravity at work are easier to understand in Newtonian terms than boomeranging. Nor do I know any easier doorway to Einstein's concept of gravity as manifestation of spacetime curvature" [10]. But nowhere in Wheeler's book is there any discussion of *empirical evidence* for "boomeranging". No doubt, Newton, Einstein and Wheeler would all have been delighted to see the simple harmonic motion demonstrated as a laboratory experiment.

2 Feasibility

Since the predicted effect has never been observed at all, our initial goal should simply be to ascertain that the oscillation prediction is a correct approximation. After laying out a basic strategy for doing the experiment, this paper concludes with a few additional remarks concerning motivation.

Apparatus that would have sufficed for our purpose were considered in the 1960s–1970s to measure G. Y. T. Chen discusses these through-the-center oscillation devices in his 1988 review paper on G measurements [11]. Each example in this group of proposals was intended for space-borne satellite laboratories. The original motivation for these ideas was to devise ways to improve the accuracy of our knowledge of G by timing the oscillation period of the simple harmonic motion. Though having some advantages over Earth-based G measurements, they also had drawbacks which ultimately prohibited them from ever being carried out.

What distinguishes these proposals from experiments that have actually been carried out in Earth-based laboratories is that the test objects were to be allowed to fall freely back and forth between extremities inside a source mass the whole time. Whereas G measurements conducted on Earth typically involve restricting the test mass's movement and measuring the force needed to do so. The most common, and historically original, method for doing this is to use a torsion balance in which a fiber provides a predetermined resistance to rotation. Torsion balances have also been used to test Einstein's Equivalence Principle (e.g. Gundlach et al. [12]). Another distinguishing characteristic of Earth-based G measurements and Equivalence Principle tests is that the test masses typically remain outside the larger source masses. Since movement of the test masses is restricted to a small range of motion, these tests can be characterized as static measurements. Torsion balance experiments in which the test mass is inside the source mass have also been performed (for example, Spero

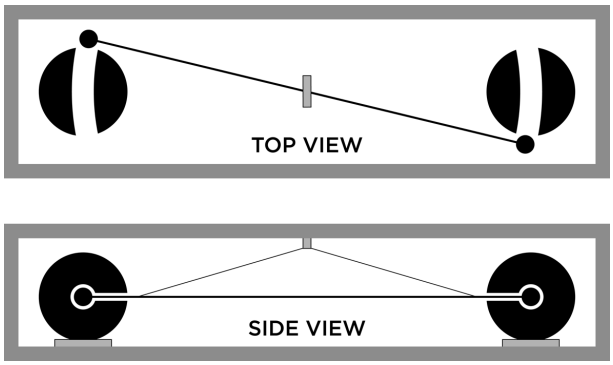


Fig. 1: Schematic of modified Cavendish balance. Since the idea is to demonstrate the simple harmonic motion only as a first approximation, deviation due to the slight arc in the trajectory is inconsequential.

et al. [13] and Hoskins et al. [14]). These latter experiments were tests of the inverse square law.

All three of these types of experiments — G measurements, Equivalence Principle and inverse-square law tests — however, are *static* measurements in the sense that the test masses were not free to move beyond a small distance compared to the size of the source mass. The key innovation in the present proposal is that we want to see an object fall radially as long as it will; we want to eliminate (ideally) or minimize (practically) any obstacle to the radial free-fall trajectory. Space-based experiments would clearly be the optimal way to achieve this. But a reasonably close approximation can be achieved with a modified Cavendish balance in an Earth-based laboratory.

As implied above, the key is to design a suspension system which, instead of providing a restoring force that prevents the test masses from moving very far, allows unrestricted or nearly unrestricted movement. Two available possibilities are fluid suspensions and magnetic suspensions (or a combination of these). In 1976 Faller and Koldewyn succeeded in using a magnetic suspension system to get a G measurement [15, 16]. The experiment's accuracy was not an improvement over that gotten by other methods, but was within 1.5% of the standard value.

As Michelini pointed out in his missing vacuum range discussion, in most G measurements the source masses are located outside an enclosure. Even in the apparatus Cavendish used for his original G measurement at atmospheric pressure, the torsion arm and test masses were isolated from the source masses by a wooden box. In Faller and Koldewyn's experiment the arm was isolated from the source masses by a vacuum chamber. The modified design requires that there be no such isolation, as the arm needs to swing freely through the center of the source masses (see Fig. 1). Given the modest goal of the present proposal, it is reasonable to expect that the technology used by Faller and Koldewyn could be adapted to test the oscillation prediction. Moreover, it seems reasonable to expect that advances in technology since 1976 (e.g. better magnets, better electronics, etc.), would make the experiment quite doable for an institution grade physics laboratory.

3 Motivation: Completeness and Aesthetics

One hardly needs to mention the many successes of Newtonian gravity. By success we mean, of course, that empirical observations match the theoretical predictions. Einsteinian gravity is even more successful. The purpose of many contemporary gravity experiments is to detect physical manifestations of the differences between Newton's and Einstein's theories. In every case Einstein's theory has proven to be more accurate. This is impressive. Given the level of thoroughness and sophistication in gravity experimentation these days, one may be taken aback to realize that Newton's and Einstein's theories *both remain untested* with regard to the problem discussed above. The simple harmonic motion prediction is so common and so obvious that we have come to take it for granted. When discussing the *prediction* for this basic experiment in weak field gravity, it would surely be more satisfactory if we could at the same time cite the *physical evidence*.

The Newtonian explanation for the predicted harmonic motion is that a massive sphere produces a force (or potential) of gravitational attraction. The corresponding general relativistic explanation is that the curvature of spacetime causes the motion. Specifically, the predicted effect is due to the slowing of clock rates toward the center of the sphere. A physical demonstration of the effect would thus indirectly, though convincingly, support general relativity's prediction that the rate of a clock at the body's center is a local minimum — a prediction that has otherwise not yet been confirmed.

In summary, if R represents the surface of a spherical mass, our empirical knowledge of how things move because of the mass within R is essentially confined to the region, $r \gtrsim R$. The region $0 \leq r \lesssim R$ is a rather fundamental and a rather large gap. It is clearly the most ponderable part of the domain. Why not fill this gap?

One of the distinctive features of the kind of experiment proposed above is that its result is, in principle, independent of size. The satellite versions mentioned by Chen were thus referred to as "clock mode" experiments. The determining factor in the oscillation period is the density of the source mass. If the source mass is made of lead (density, $\rho \approx 11,000$ kg/m^3) the oscillation period is about one hour. Would it not be fascinating to observe for an hour, to watch the oscillation take place, knowing that the mass of the larger body is the essential thing making it happen? In my opinion this would be a beautiful sight. Beautiful for completing the domain, $0 \leq r \lesssim R$, and beautiful simply to see what no human being has seen before.

Submitted on September 09, 2010 / Accepted on September 15, 2010

References

1. Michelini M. The Missing Measurements of the Gravitational Constant. *Progress in Physics*, 2009, v. 3, 64–68.
2. Halliday D., Resnick R., Walker J. Fundamentals of Physics. Wiley, New York, 1993.
3. Valens E. G. The Attractive Universe: Gravity and the Shape of Space. World, Cleveland, 1969.
4. Tipler P. A. Physics. Worth, New York, 1982.
5. French A. P. Newtonian Mechanics. Norton, New York, 1971.
6. Misner C. W., Thorne K. and Wheeler J. A. Gravitation. W. H. Freeman, San Francisco, 1973.
7. Tangherlini F. R. An Introduction to the General Theory of Relativity. *Nuovo Cimento Supplement*, 1966, v. 20, 1–86.
8. Epstein L. C. Relativity Visualized. Insight, San Francisco, 1988, pp. 157–158.
9. Taylor N. W. Note on the Harmonic Oscillator in General Relativity. *Australian Journal of Mathematics*, 1961, v. 2, 206–208.
10. Wheeler J. A. A Journey Into Gravity and Spacetime. Scientific American Library, New York, 1990.
11. Chen Y. T. The Measurement of the Gravitational Constant. *International Symposium on Experimental Gravitational Physics*, edited by Michelson P. F., Enke H., Pizzella G. World Scientific, Singapore, 1988, pp. 90–109.
12. Gundlach J. H., Smith G. L., Adelberger E. G., Heckel B. R., Swanson H. E. Short-Range Test of the Equivalence Principle. *Physical Review Letters*, 1997, v. 78, 2523–2526.
13. Spero R. E., Hoskins J. K., Newman R., Pellam J., Schultz J. Test of the Gravitational Inverse-Square Law at Laboratory Distances. *Physical Review Letters*, 1980, v. 44, 1645–1648.
14. Hoskins J. K., Newman R. D., Spero R. E., Schultz J. Experimental tests of the gravitational inverse-square law for mass separations from 2 to 105 cm. *Physical Review D*, 1985, v. 32, 3084–3095.
15. Koldewyn W. A. A New Method for Measuring the Newtonian Gravitational Constant, G. PhD thesis, Wesleyan University, 1976.
16. Faller J. E. A Prototype Measurement of the Newtonian Gravitational Constant Using an Active Magnetic Suspension Torsion Fiber. *Proceedings of the 1983 International School and Symposium on Precision Measurement and Gravity Experiment*, edited by Wei-Tou N., National Tsing Hua University, Hsinchu, Taiwan, 1983, pp. 541–556.

Experimental Investigation of the Fresnel Drag Effect in RF Coaxial Cables

Reginald T. Cahill* and David Brotherton†

*School of Chemical and Physical Sciences, Flinders University, Adelaide 5001, Australia.
E-mail: Reg.Cahill@flinders.edu.au

†Geola Technologies Ltd, Sussex Innovation Centre, Science Park Square, Falmer, East Sussex, BN1 9SB, United Kingdom.
E-mail: dbr@geola.co.uk

An experiment that confirms the Fresnel drag formalism in RF coaxial cables is reported. The Fresnel 'drag' in bulk dielectrics and in optical fibers has previously been well established. An explanation for this formalism is given, and it is shown that there is no actual drag phenomenon, rather that the Fresnel drag effect is merely the consequence of a simplified description of EM scattering within a dielectric in motion wrt the dynamical 3-space. The Fresnel drag effect plays a critical role in the design of various light-speed anisotropy detectors.

1 Introduction

In 2002 it was discovered that the Michelson-Morley 1887 light-speed anisotropy experiment [1], using the interferometer in gas mode, had indeed detected anisotropy, by taking account of both a physical Lorentz length contraction effect for the interferometer arms, and the refractive index effect of the air in the light paths [2, 3]. The observed fringe shifts corresponded to an anisotropy speed in excess of 300 km/s. While confirmed by numerous later experiments, particularly that of Miller [4], see [6] for an overview, the most accurate analysis used the Doppler shifts from spacecraft earth-flybys [5, 6], which gave the solar-system a galactic average speed through 3-space of 486 km/s in the direction $RA = 4.29^h$, $Dec = -75.0°$, a direction within 5° of that found by Miller in his 1925/26 gas-mode Michelson interferometer experiment*. In vacuum mode a Michelson interferometer cannot detect the anisotropy, nor its turbulence effects, as shown by the experiments in [7–12], actually using resonant orthogonal cavities. These experiments show, overall, the difference between Lorentzian Relativity (LR) and Special Relativity (SR). In LR the length contraction effect is caused by motion of a rod, say, through the dynamical 3-space, whereas in SR the length contraction is only a perspective effect, occurring only when the rod is moving relative to an observer. This was further clarified when an exact mapping between Galilean space and time coordinates and the Minkowski-Einstein spacetime coordinates was recently discovered [13]. This demonstrates that the SR time dilation and space contraction effects are merely the result of using an unphysical choice of space and time coordinates that, by construction, makes the speed of light in vacuum an invariant, but only wrt to that choice of coordinates. Such a contrived invariance has no connection with whether light speed anisotropy is detectable or not — that is to be determined by experiments.

*This speed and direction is very different to the CMB speed and direction — which is an unrelated phenomenon.

The detection of light speed anisotropy — revealing a flow of space past the detector, is now entering an era of precision measurements. These are particularly important because experiments have shown large turbulence effects in the flow, and are beginning to characterise this turbulence. Such turbulence can be shown to correspond to what are, conventionally, known as gravitational waves, although not those implied by General Relativity, as they are much larger than these [14–16].

The detection and characterisation of these wave/ turbulence effects requires the development of small and cheap detectors, such as optical fiber Michelson interferometers [18]. However in all detectors the EM signals travel through a dielectric, either in bulk or optical fiber or through RF coaxial cables. For this reason it is important to understand the so-called Fresnel drag effect. In optical fibers the Fresnel drag effect has been established [17]. This is important in the operation of Sagnac optical fiber gyroscopes, for then the calibration is independent of the fiber refractive index. The Fresnel drag speed is a phenomenological formalism that characterises the effect of the absolute motion of the propagation medium upon the speed of the EM radiation within that medium.

The Fresnel drag expression is that a dielectric in absolute motion through space at speed v causes the EM radiation to travel at speed

$$V(v) = \frac{c}{n} + v\left(1 - \frac{1}{n^2}\right) \qquad (1)$$

wrt the dielectric, when V and v have the same direction. Here n is the dielectric refractive index. The 2nd term is known as the Fresnel drag, appearing to show that the moving dielectric "drags" the EM radiation, although below we show that this is a misleading interpretation. That something unusual was happening followed from the discovery of stellar aberration by Bradley in 1725. Here the direction of the telescope must be varied over a year when observing a given star. This is caused by the earth's orbital speed of 30 km/s. Then Airy

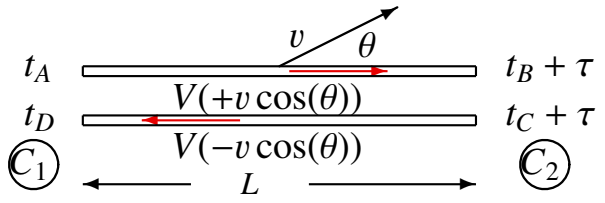


Fig. 1: Schematic layout for measuring the one-way speed of light in either free-space, optical fibers or RF coaxial cables, without requiring the synchronisation of the clocks C_1 and C_2. Here τ is the unknown offset time between the clocks, and $t_A, t_B + \tau, t_C + \tau, t_D$ are the observed clock times, while t_B, t_C are, *a priori*, unknown true times. V is the light speed in (1), and v is the speed of the apparatus through space, in direction θ.

in 1871 demonstrated that the same aberration angle occurs even when the telescope is filled with water. This effect is explained by the Fresnel expression in (1), which was also confirmed by the Fizeau experiment in 1851, who used two beams of light travelling through two tubes filled with flowing water, with one beam flowing in the direction of the water, and the other counterflowing. Interferometric means permitted the measurement of the travel time difference between the two beams, confirming (1), with v the speed of the water flow relative to the apparatus. This arrangement cannot detect the absolute motion of the solar system, as this contribution to the travel time difference cancels because of the geometry of the apparatus.

There have been various spurious "derivations" of (1), some attempting to construct a physical "drag" mechanism, while another uses the SR addition formula for speeds. However that well-known addition formula is merely a mathematical manifestation of using the unphysical Minkowski-Einstein coordinates noted above, and so is nothing but a coordinate effect, unrelated to experiment. Below we give a simple heuristic derivation which shows that there is no actual "drag" phenomenon. But first we show the unusual consequences of (1) in one-way speed of EM radiation experiments. It also plays a role in 2nd order v/c experiments, such as the optical-fiber Michelson interferometer [18].

2 One-way Speed of Light Anisotropy Measurements

Fig.1 shows the arrangement for measuring the one-way speed of light, either in vacuum, a dielectric, or RF coaxial cable. It is usually argued that one-way speed of light measurements are not possible because the clocks cannot be synchronised. Here we show that this is false, and at the same time show an important consequence of (1). In the upper part of Fig.1 the actual travel time t_{AB} from A to B is determined by

$$V(v\cos(\theta))t_{AB} = |\mathbf{L}'| \qquad (2)$$

where $|\mathbf{L}'| = |\mathbf{L} + \mathbf{v}t_{AB}| \approx L + v\cos(\theta)t_{AB} + ..$ is the actual distance travelled, at speed $V(v\cos(\theta))$, using $vt_{AB} \ll L$, giving

$$V(v\cos(\theta))t_{AB} = L + v\cos(\theta)t_{AB} + ... \qquad (3)$$

where the 2nd term comes from, approximately, the end B moving an additional distance $v\cos(\theta)t_{AB}$ during the true time interval t_{AB}. This gives

$$t_{AB} \approx \frac{L}{V(v\cos(\theta)) - v\cos(\theta)} = \frac{nL}{c} + \frac{v\cos(\theta)L}{c^2} + .. \qquad (4)$$

on using (1) and expanding to 1st oder in v/c. If we ignore the Fresnel drag term in (1) we obtain, instead,

$$t_{AB} \approx \frac{L}{c/n - v\cos(\theta)} = \frac{nL}{c} + \frac{n^2 v\cos(\theta)L}{c^2} + .. \qquad (5)$$

The 1st important observation is that the v/c component in (4) is independent of the dielectric refractive index n. This is explained in the next section. If the clocks were synchronised t_{AB} would be known, and by changing direction of the light path, that is varying θ, the magnitude of the 2nd term may be separated from the magnitude of the 1st term. If the clocks are not synchronised, then the measured travel time $\bar{t}_{AB} = (t_B + \tau) - t_A = t_{AB} + \tau$, where τ is the unknown, but fixed, offset between the two clocks. But this does not change the above argument. The magnitude of v and the direction of v can be determined by varying θ. For a small detector the change in θ can be achieved by a direct rotation. But for a large detector, such as De Witte's [19] 1.5 km RF coaxial cable experiment, the rotation was achieved by that of the earth. The reason for using opposing propagation directions, as in Fig.1, and then measuring travel time differences, is that local temperature effects cancel. This is because a common temperature change in the two adjacent cables changes the speed to the same extent, whereas absolute motion effects cause opposite signed speed changes. Then the temperature effects cancel on measuring differences in the travel times, whereas absolute motion effects are additive. Finally, after the absolute motion velocity has been determined, the two spatially separated clocks may be synchronised.

That the v/c term in t_{AB} in (4) is independent of n means that various techniques to do a 1st order in v/c experiment that involves using two dielectrics with different values of n fail. One such experiment was by Trimmer et al [20], who used a triangular interferometer, with the light path split into one beam passing through vacuum, and the other passing through glass. No 1st order effect was seen. This is because the v-dependent travel times through the glass, and corresponding vacuum distance, have the same value to 1st order in v/c. On realising this Trimmer et al. subsequently withdrew their paper, see reference [21]. Cahill [22] performed a dual optical-fiber/RF coaxial cable experiment that was supposedly 1st order in v/c. If the Fresnel drag formalism applies to both optical fibers and RF coaxial cables, then again there could not

have been any v/c signal in that experiment, and the observed effects must have been induced by temperature effects. All this implies that because of the Fresnel drag effect it appears not possible to perform a v/c experiment using one clock — rather two must be used, as in Fig.1. This, as noted above, does not require clock synchronisation, but it does require clocks that very stable. To use one "clock" appears then to require 2nd order in v/c detectors, but then the effect is some 1000 times smaller, and requires interferometric methods to measure the very small travel time differences, as in gas-mode and optical-fiber Michelson interferometers. It is indeed fortuitous that the early experiments by Michelson and Morley, and by Miller, were in gas mode, but not by design.

The Krisher optical fiber 1st order v/c experiment [23] measured the phase differences ϕ_1 and ϕ_2 between the two signals travelling in different directions through very long optical fibers, rather than the travel time variations, as the earth rotated. This involves two phase comparators, with one at each end of the fibers. However the phases always have a multiple of 2π phase ambiguity, and in the Krisher experiment this was overlooked. However the timing of the maxima/minima permitted the Right Ascension (RA) of the direction of v to be determined, as the direction of propagation is changed by rotation, and the result agreed with that found by Miller; see [6] for plots of the Krisher data plotted against local sidereal times.

3 Deriving the Fresnel Drag Formalism

Here we give a heuristic derivation of the Fresnel drag speed formalism in a moving dielectric, with the dielectric modeled by random geometrical-optics paths, see Fig.2. These may be thought of as modelling EM wave scatterings, and their associated time delays. The slab of dielectric has length L and travels through space with velocity $\mathbf{v}$, and with EM radiation traveling, overall, from A to B. The top of Fig.2 shows the microscopic heuristic model of propagation through the dielectric with EM radiation traveling at speed c wrt space between scattering events, being scattered from random sites — atoms, moving through space with velocity $\mathbf{v}$. The bottom of Fig.2 shows the macroscopic description with EM radiation effectively traveling in a straight line directly from A to B, with effective linear speed $V(v\cos(\theta))$, and with the dielectric now described by a refractive index n.

The key insight is that when a dielectric has absolute velocity $\mathbf{v}$ through space, the EM radiation travels at speed c wrt space, between two scattering events within the dielectric. Consider a straight line propagation between scattering events e and f, with angle ϕ to $\mathbf{v}$, see Fig.2. Consider the paths from the rest frame of the space. The EM wave must travel to a point in space f', and then the distance travelled dl', at speed c, is determined by the vector sum $\mathbf{dl}' = \mathbf{dl} + \mathbf{v}dt$, with dl the distance between scattering points e and f, defined in the rest frame of the matter, and $\mathbf{v}dt$ is the displacement of f

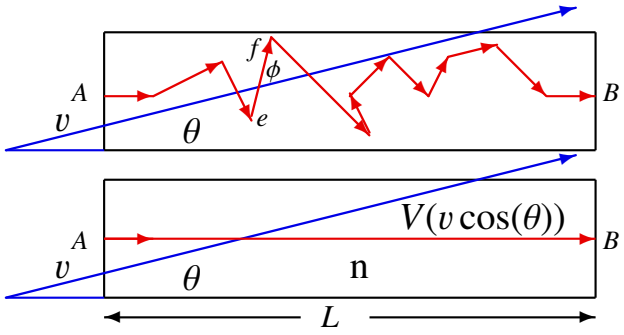


Fig. 2: Slab of dielectric, length L, traveling through space with velocity $\mathbf{v}$, and with EM radiation traveling, overall, from A to B, drawn in rest frame of slab. Top: Microscopic model showing scattering events, with free propagation at speed c relative to the space, between scattering events. Bottom: The derived macroscopic phenomenological description showing the signal travelling at speed $V(v\cos(\theta))$, as given by the Fresnel drag expression in (1). The dielectric refractive index is n.

to f', because of the absolute motion of the scattering atoms. Then the travel time to 1st order in v/c is

$$dt = \frac{dl'}{c} \approx \frac{dl}{c} + \frac{v\cos(\phi)dt}{c} + ..., \quad \text{giving} \tag{6}$$

$$dt \approx \frac{dl}{c} + \frac{vdl\cos(\phi)}{c^2} + ... = \frac{dl}{c} + \frac{\mathbf{v}\cdot\mathbf{dl}}{c^2} + ... \tag{7}$$

We ignore Lorentz length contraction of the slab as this only contributes at 2nd order in v/c. Summing over paths to get total travel time from A to B

$$\begin{aligned} t_{AB} &= \int_A^B \frac{dl}{c} + \int_A^B \frac{\mathbf{v}\cdot\mathbf{dl}}{c^2} + ... \\ &= \frac{l}{c} + \frac{Lv\cos(\theta)}{c^2} + ... \\ &= \frac{nL}{c} + \frac{Lv\cos(\theta)}{c^2} + ..., \end{aligned} \tag{8}$$

where L is the straight line distance from A to B in the matter rest frame, and $n = l/L$ defines the refractive index of the dielectric in this treatment, as when the dielectric is at rest the effective speed of EM radiation through matter in a straight line from A to B is defined to be c/n. Note that t_{AB} does not involve n in the v dependent 2nd term. This effect is actually the reason for the Fresnel drag formalism. The macroscopic treatment, which leads to the Fresnel drag formalism, involves the sum $|\mathbf{L}'| = |\mathbf{L} + \mathbf{v}t_{AB}|$, for the macrosocopic distance traveled, which gives for the travel time

$$\begin{aligned} t_{AB} &= \frac{L'}{V} \approx \frac{L}{V(v\cos(\theta))} + \frac{v\cos(\theta)t_{AB}}{V(v\cos(\theta))}, \quad \text{giving} \\ t_{AB} &= \frac{L}{V(v\cos(\theta))} + \frac{Lv\cos(\theta)}{V(v\cos(\theta))^2} + ... \end{aligned} \tag{9}$$

where $V(v\cos(\theta))$ is the effective linear speed of EM radiation in direction AB at angle θ to $\mathbf{v}$, and $v\cos(\theta)t_{AB}$ is the

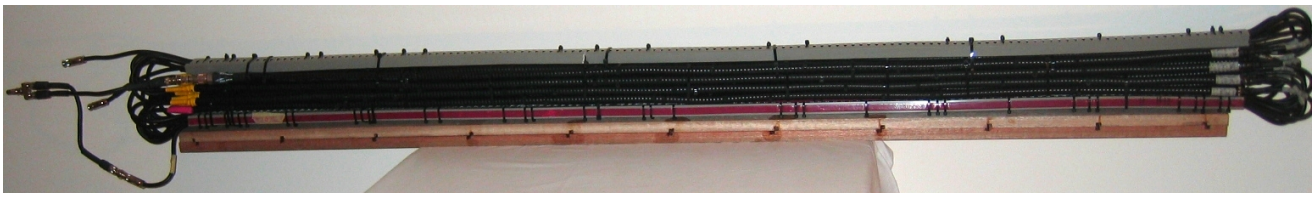


Fig. 3: Photograph of the RF coaxial cables arrangement, based upon 16×1.85 m lengths of phase stabilised Andrew HJ4-50 coaxial cable. These are joined to 16 lengths of FSJ1-50A cable, in the manner shown schematically in Fig.4. The 16 HJ4-50 coaxial cables have been tightly bound into a 4×4 array, so that the cables, locally, have the same temperature, with cables in one of the circuits embedded between cables in the 2nd circuit. This arrangement of the cables permits the cancellation of temperature effects in the cables. A similar array of the smaller diameter FSJ1-50A cables is located inside the grey-coloured conduit boxes. This arrangement has permitted the study of the Fresnel drag effect in RF coaxial cables, and revealed that the usual Fresnel drag speed expression applies.

extra distance travelled, caused by the end B moving. This form assumes that the total distance L' is travelled at speed $V(v\cos(\theta))$. This reproduces the microscopic result (8) only if $V(v) = c/n + v(1 - 1/n^2)$, which is the Fresnel drag expression. The key point is that the Fresnel drag formalism is needed to ensure, despite appearances, that the extra distance traveled due to the absolute motion of the dielectric, is travelled at speed c, and not at speed c/n, even though the propagation is within the dielectric. Hence there is no actual drag phenomenon involved, and so the nomenclature "Fresnel drag" is misleading.

However it was not clear that the same analysis applied to RF coaxial cables, because of the possible effects of the conduction electrons in the inner and outer conductors. The dual coaxial cable experiment reported herein shows that the Fresnel drag expression also applies in this case. The Fresnel drag effect is a direct consequence of the absolute motion of the slab of matter through space, with the speed of EM radiation being c wrt space itself. A more complete derivation based on the Maxwell-Hertz equations is given in Drezet [24].

4 Fresnel Drag Experiment in RF Coaxial Cables

We now come to the 1st experiment that has studied the Fresnel drag effect in RF coaxial cables. This is important for any proposed EM anisotropy experiment using RF coaxial cables. The query here is whether the presence of the conductors forming the coaxial cables affects the usual Fresnel drag expression in (1), for a coaxial cable has an inner and outer conductor, with a dielectric in between.

Fig.4 shows the schematic arrangement using two different RF coaxial cables, with two separate circuits, and Fig.3 a photograph. One measures the travel time difference of two RF signals from a Rubidium frequency standard (Rb) with a Digital Storage Oscilloscope (DSO). In each circuit the RF signal travels one-way in one type of coaxial cable, and returns via a different kind of coaxial cable. Two circuits are used so that temperature effects cancel — if a temperature change alters the speed in one type of cable, and so the travel time, that travel time change is the same in both circuits, and

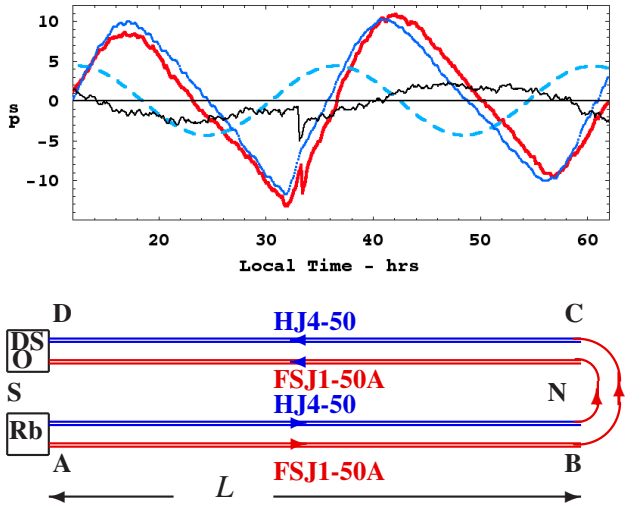


Fig. 4: Top: Data, from May17-19, 2010, from the dual RF coaxial cable experiment enabling Fresnel drag in coaxial cables to be studied: Red plot is relative 10 MHz RF travel times between the two circuits, and blue plot is temperature of the air (varying from 19 to 23°C) passing into the LeCroy DSO, scaled to fit the travel time data. The black plot is travel time differences after correcting for DSO temperature effects. The dashed plot is time variation expected using spacecraft earth-flyby Doppler shift determined velocity, if the Fresnel drag effect is absent in RF coaxial cables. Bottom: Schematic layout of the coaxial cables. This ensures two opposing circuits that enable cancellation of local temperature effects in the cables. In practice the cables are divided further, as shown in Fig.3.

cancels in the difference. Even though phase-stabilised coaxial cables are used, the temperature effects need to be cancelled in order to be able to reliably measure time differences at ps levels. To ensure cancellation of temperature effects, and also for practical convenience, the Andrew HJ4-50 cables are cut into 8×1.85 m shorter lengths in each circuit, corresponding to a net length of $L = 8\times1.85 = 14.8$ m. The curved parts of the Andrew FSJ1-50A cables contribute only at 2nd order in v/c.

To analyse the experimental data we modify the Fresnel

drag speed expression in (1) to

$$V(v) = \frac{c}{n_i} + v\left(1 - \frac{1}{m_i^2}\right) \tag{10}$$

for each cable, $i = 1, 2$, where $m_i = n_i$ gives the normal Fresnel drag, while $m_i = 1$ corresponds to no Fresnel drag. Repeating the derivation leading to (4) we obtain to 1st order in v/c the travel time difference between the two circuits,

$$\Delta t = \frac{2Lv\cos(\theta)}{c^2}\left(\frac{n_1^2}{m_1^2} - \frac{n_2^2}{m_2^2}\right) \tag{11}$$

The apparatus was orientated NS and used the rotation of the earth to change the angle θ. Then θ varies between $\lambda + \delta - 90^\circ = 20^\circ$ and $\lambda - \delta + 90^\circ = 50^\circ$, where $\lambda = 35^\circ$ is the latitude of Adelaide, and $\delta = 75^\circ$ is the declination of the 3-space flow from the flyby Doppler shift analysis, and with a speed of 486 km/s.. Then if $m_i \neq n_i$ a signal with period 24 h should be revealed. We need to compute the magnitude of the time difference signal if there is no Fresnel drag effect. The FSJ1-50A has an effective refractive index $n_1 = 1.19$, while the HJ4-50 has $n_2 = 1.11$, and then Δt would change by 8.7 ps over 24 hours, and have the phase shown in Fig.4. However while cable temperature effects have been removed by the cable layout, another source of temperature effects is from the LeCroy WaveRunner WR6051A DSO. To achieve ps timing accuracy and stability the DSO was operated in Random Interleaved Sampling (RIS) mode. This uses many signal samples to achieve higher precision. However in this mode the DSO temperature compensation re-calibration facility is disabled. To correct for this it was discovered that the timing errors between the two DSO channels very accurately tracked the temperature of the cooling air being drawn into the DSO. Hence during the experiment that air temperature was recorded. The Rb frequency standard was a Stanford Research Systems FS725. The results for 48 hours in mid May, 2010, are shown in Fig. 4: The red plot, with glitch, shows the DSO measured time difference values, while the blue plot shows the temperature variation of the DSO air-intake temperature, scaled to the time data. We see that the time data very closely tracks the air-intake temperature. Subtracting this temperature effect we obtain the smaller plot, which has a range of 5 ps, but showing no 24 h period. The corrected timing data may still have some small temperature effects. The glitch in the timing data near local time of 34 h was probably caused by a mechanical stress-release event in the cables. Hence the data implies that there is no 1st order effect in v/c, and so, from (11), that $n_1/m_1 = n_2/m_2$, with the simplest interpretation being that, in each cable $m = n$. This means that the Fresnel drag effect expression in (1) applies to RF coaxial cables.

5 Conclusions

The first experiment to study the Fresnel drag effect in RF coaxial cables has revealed that these cables exhibit the same effect as seen in bulk dielectrics and in optical fibers, and so this effect is very general, and in the case of the RF coaxial cables, is not affected by the conductors integral to RF coaxial cables. Because this experiment is a null experiment, after correcting for temperature effects in the DSO, its implications follow only when the results are compared with non-null experiments. Here we have compared the results with those from the spacecraft earth-flyby Doppler shift data results. Then we can deduce that the null result is caused by the Fresnel drag effect in the cables, and not by the absence of light speed anisotropy. This is to be understood from the heuristic derivation given herein, where it was shown that the Fresnel drag expression actually involves no drag effect at all, rather its form is such as to ensure that between scatterings the EM waves travel at speed c wrt to the 3-space, that is, that the "speed of light" is not an invariant. This experiment, as have many others, shows that the speed of light, as measured by an observer, actually depends on the speed of that observer wrt to 3-space. We have also shown how the speed of light may be measured in a one-way 1st order in v/c experiment, using spatially separated clocks that are not *a priori* synchronised, by rotating the apparatus. Subsequently, once the velocity of space past the detector is known, the clocks may be synchronised by light speed signalling.

Acknowledgements

Special thanks to Professor Igor Bray for supporting this research.

Submitted on August 23, 2010 / Accepted on September 20, 2010

References

1. Michelson A. A., Morley E. W. On the Relative Motion of the Earth and the Luminiferous Ether, *American Journal of Science*, 1887, v. 34, 333–345.
2. Cahill R. T., Kitto K. Michelson-Morley Experiments Revisited. *Apeiron*, 2003, v. 10, no. 2, 104–117.
3. Cahill R. T. The Michelson and Morley 1887 Experiment and the Discovery of Absolute Motion. *Progress in Physics*, 2005, v. 3, 25–29.
4. Miller D. C. The Ether-Drift Experiments and the Determination of the Absolute Motion of the Earth. *Review of Modern Physics*, 1933, v. 5, 203–242.
5. Anderson J. D., Campbell J. K., Ekelund J. E., Ellis J., Jordan J. F. Anomalous Orbital-Energy Changes Observed during Spaceraft Flybys of Earth. *Physical Review Letters*, 2008, v. 100, 091102.
6. Cahill R. T. Combining NASA/JPL One-Way Optical-Fiber Light-Speed Data with Spacecraft Earth-Flyby Doppler-Shift Data to Characterise 3-Space Flow. *Progress in Physics*, 2009, v. 4, 50–64.
7. Braxmaier C., Müller H., Pradl O., Mlynek J., Peters A., Schiller S. Tests of Relativity Using a Cryogenic Optical Resonator. *Physical Review Letters*, 2002, v. 88, 010401.
8. Müller H., Braxmaier C., Herrmann S., Peters A., Lämmerzahl C. Electromagnetic cavities and Lorentz invariance violation. *Physical Review D*, 2003, v. 67, 056006.
9. Wolf P., Bize S., Clairon A., Santarelli G., Tobar M. E., Luiten A. N. Improved test of Lorentz invariance in electrodynamics. *Physical Review D*, 2004, v. 70, 051902.

10. Wolf P., Bize S., Clairon A., Luiten A. N., Santarelli G., Tobar M. E. Tests of Lorentz Invariance using a Microwave Resonator. *Physical Review Letters*, v. 90, no. 6, 060402.

11. Lipa J. A., Nissen J. A., Wang S., Stricker D. A., Avaloff D. New Limit on Signals of Lorentz Violation in Electrodynamics. *Physical Review Letters*, 2003, v. 90, 060403.

12. Eisle Ch., Nevsky A. Yu., Schiller S. Laboratory Test of the Isotropy of Light Propagation at the 10^{-17} Level. *Physical Review Letters*, 2009, v. 103, 090401.

13. Cahill R. T. Unravelling Lorentz Covariance and the Spacetime Formalism. *Progress in Physics*, 2008, v. 4, 19–24.

14. Cahill R. T. Process Physics: From Information Theory to Quantum Space and Matter. Nova Science Pub., New York, 2005.

15. Cahill R. T. Dynamical 3-Space: A Review, in: Ether space-time and cosmology: New insights into a key physical medium, Duffy M. and Lévy J. (Editors), Apeiron, 2009, pp. 135-200.

16. Cahill R. T. Quantum Foam, Gravity and Gravitational Waves, in: Relativity, Gravitation, Cosmology: New Developments, Dvoeglazov V. (Editor), Nova Science Pub., New York, 2010, pp. 1-55.

17. Wang R., Zhengb Yi., Yaob A., Langley D. Modified Sagnac Experiment for Measuring Travel-time Difference Between Counterpropagating Light Beams in a Uniformly Moving Fiber. *Physics Letters A*, 2003, no. 1-2, 7–10.

18. Cahill R. T., Stokes F. Correlated Detection of sub-mHz Gravitational Waves by Two Optical-Fiber Interferometers. *Progress in Physics*, 2008, v. 2, 103–110.

19. Cahill R. T. The Roland De Witte 1991 Experiment. *Progress in Physics*, 2006, v. 3, 60–65.

20. Trimmer W. S. N. Baierlein R. F., Faller J. E., Hill H. A. Experimental Search for Anisotropy in the Speed of Light. *Physical Review D*, 1973, v. 8, 3321–3326.

21. Trimmer W. S. N. Baierlein R. F., Faller J. E., Hill H. A. Erratum: Experimental Search for Anisotropy in the Speed of Light. *Physical Review D*, 1974, v. 9, 2489.

22. Cahill R. T. A New Light-Speed Anisotropy Experiment: Absolute Motion and Gravitational Waves Detected. *Progress in Physics*, 2006, v. 4, 73–92.

23. Krisher T. P., Maleki L., Lutes G. F., Primas L. E., Logan R. T., Anderson J. D., Will C. M. Test of the Isotropy of the One-Way Speed of Light using Hydrogen-Maser Frequency Standards. *Physical Review D*, 1990, v. 42, 731–734.

24. Drezet A. The Physical Origin of the Fresnel Drag of Light by a Moving Dielectric Medium. *European Physics Journal B*, 2005, v. 45, no. 1, 103–110.

Dynamical 3-Space Gravity Theory: Effects on Polytropic Solar Models

Richard D. May and Reginald T. Cahill

School of Chemical and Physical Sciences, Flinders University, Adelaide 5001, Australia
E-mail: Richard.May@flinders.edu.au, Reg.Cahill@flinders.edu.au

Numerous experiments and observations have confirmed the existence of a dynamical 3-space, detectable directly by light-speed anisotropy experiments, and indirectly by means of novel gravitational effects, such as bore hole g anomalies, predictable black hole masses, flat spiral-galaxy rotation curves, and the expansion of the universe, all without dark matter and dark energy. The dynamics for this 3-space follows from a unique generalisation of Newtonian gravity, once that is cast into a velocity formalism. This new theory of gravity is applied to the solar model of the sun to compute new density, pressure and temperature profiles, using polytrope modelling of the equation of state for the matter. These results should be applied to a re-analysis of solar neutrino production, and to stellar evolution in general.

1 Introduction

It has been discovered that Newton's theory of gravity [1] missed a significant dynamical process, and a uniquely determined generalisation to include this process has resulted in the explanation of numerous gravitational anomalies, such as bore hole g anomalies, predictable black hole masses, flat spiral-galaxy rotation curves, and the expansion of the universe, all without dark matter and dark energy [2–4]. This theory of gravity arises from the dynamical 3-space, described by a dynamical velocity field, when the Schrödinger equation is generalised to take account of the propagation of quantum matter in the dynamical 3-space. So gravity is now an emergent phenomenon, together with the equivalence principle.

The dynamical 3-space has been directly observed using various light-speed anisotropy experiments, dating from the 1st detection by Michelson and Morley in 1887 [5,6], giving a speed in excess of 300 km/s, after re-calibrating the gas-mode interferometer for actual length contraction effects, to the latest using spacecraft earth-flyby Doppler shift data [7]. Overall these experiments reveal that relativistic effects are caused by the absolute motion of rods and clocks wrt the dynamical 3-space, essentially Lorentzian Relativity (LR), rather than the Special Relativity (SR) formalism, which has recently been shown by means of an exact change of space and time variables, to be equivalent to Galilean Relativity [8].

Here we apply the new gravity theory to the internal dynamics of the sun, and compute new density, pressure and temperature profiles, using the polytrope model for the equation of state of the matter. These results should then be applied to a re-analysis of neutrino production [9]. In general the Newtonian-gravity based standard model of stellar evolution also needs re-examination.

2 Dynamical 3-Space

Newton's inverse square law of gravity has the differential form

$$\nabla \cdot \mathbf{g} = -4\pi G\rho, \quad \nabla \times \mathbf{g} = \mathbf{0}, \tag{1}$$

for the acceleration field $\mathbf{g}(\mathbf{r}, t)$, assumed to be fundamental and existing within Newton's model of space, which is Euclidean, static, and unobservable. Application of this to spiral galaxies and the expanding universe has lead to many problems, including, in part, the need to invent dark energy and dark matter*. However (1) has a unique generalisation that resolves these and other problems. In terms of a velocity field $\mathbf{v}(\mathbf{r}, t)$ (1) has an equivalent form [2, 3]

$$\nabla \cdot \left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right) = -4\pi G\rho, \quad \nabla \times \mathbf{v} = \mathbf{0}, \tag{2}$$

where now

$$\mathbf{g} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}, \tag{3}$$

is the well-known Galilean covariant Euler acceleration of the substratum that has velocity $\mathbf{v}(\mathbf{r}, t)$. Because of the covariance of $\mathbf{g}$ under a change of the spatial coordinates only relative internal velocities have an ontological existence — the coordinates $\mathbf{r}$ then merely define a mathematical embedding space.

We give a brief review of the concept and mathematical formalism of a dynamical flowing 3-space, as this is often confused with the older dualistic space and aether ideas, wherein some particulate aether is located and moving through an unchanging Euclidean space — here both the space and

*The Friedmann equation for the expanding universe follow trivially from (1), as shown in [4], but then needs "dark matter" and "dark energy" to fit the cosmological data.

the aether were viewed as being ontologically real. The dynamical 3-space is different: here we have only a dynamical 3-space, which at a small scale is a quantum foam system without dimensions and described by fractal or nested homotopic mappings [2]. This quantum foam is not embedded in any space — the quantum foam is all there is, and any metric properties are intrinsic properties solely of that quantum foam. At a macroscopic level the quantum foam is described by a velocity field $\mathbf{v}(\mathbf{r}, t)$, where $\mathbf{r}$ is merely a [3]-coordinate within an embedding space. This embedding space has no ontological existence — it is merely used to (i) record that the quantum foam has, macroscopically, an effective dimension of 3, and (ii) to relate other phenomena also described by fields, at the same point in the quantum foam. The dynamics for this 3-space is easily determined by the requirement that observables be independent of the embedding choice, giving, for zero-vorticity dynamics and for a flat embedding space, and preserving the inverse square law outside of spherical masses, at least in the usual cases, such as planets,

$$\nabla\cdot\left(\frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}\right)+\frac{\alpha}{8}\left((trD)^2-tr(D^2)\right)=-4\pi G\rho,$$

$$\nabla\times\mathbf{v}=\mathbf{0},\quad D_{ij}=\frac{1}{2}\left(\frac{\partial v_i}{\partial x_j}+\frac{\partial v_j}{\partial x_i}\right),\tag{4}$$

where $\rho(\mathbf{r}, t)$ is the matter and EM energy densities, expressed as an effective matter density. Borehole g measurements and astrophysical black hole data has shown that $\alpha\approx 1/137$ is the fine structure constant to within observational errors [2,3,10]. For a quantum system with mass m the Schrödinger equation is uniquely generalised [10] with the new terms required to maintain that the motion is intrinsically wrt the 3-space, and not wrt the embedding space, and that the time evolution is unitary:

$$i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r},t)-i\hbar\left(\mathbf{v}\cdot\nabla+\frac{1}{2}\nabla\cdot\mathbf{v}\right)\psi(\mathbf{r},t).\tag{5}$$

The space and time coordinates $\{t, x, y, z\}$ in (4) and (5) ensure that the separation of a deeper and unified process into different classes of phenomena — here a dynamical 3-space (quantum foam) and a quantum matter system, is properly tracked and connected. As well the same coordinates may be used by an observer to also track the different phenomena. However it is important to realise that these coordinates have no ontological significance — they are not real. The velocities $\mathbf{v}$ have no ontological or absolute meaning relative to this coordinate system — that is in fact how one arrives at the form in (5), and so the "flow" is always relative to the internal dynamics of the 3-space. A quantum wave packet propagation analysis of (5) gives the acceleration induced by wave refraction to be [10]

$$\mathbf{g}=\frac{\partial\mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}+(\nabla\times\mathbf{v})\times\mathbf{v}_R,$$
$$\mathbf{v}_R(\mathbf{r}_o(t),t)=\mathbf{v}_o(t)-\mathbf{v}(\mathbf{r}_o(t),t),\tag{6}$$

where $\mathbf{v}_R$ is the velocity of the wave packet relative to the local 3-space, and where $\mathbf{v}_o$ and $\mathbf{r}_o$ are the velocity and position relative to the observer, and the last term in (6) generates the Lense-Thirring effect as a vorticity driven effect. Together (4) and (6) amount to the derivation of gravity as a quantum effect, explaining both the equivalence principle ($\mathbf{g}$ in (6) is independent of m) and the Lense-Thirring effect. Overall we see, on ignoring vorticity effects, that

$$\nabla\cdot\mathbf{g}=-4\pi G\rho-4\pi G\rho_{DM},\tag{7}$$

where

$$\rho_{DM}=\frac{\alpha}{32\pi G}\left((trD)^2-tr(D^2)\right).\tag{8}$$

This is Newtonian gravity but with the extra dynamical term which has been used to define an effective "dark matter" density. This is not real matter, of any form, but is the matter density needed within Newtonian gravity to explain the flat rotation curves of spiral galaxies, large light bending and lensing effects from galaxies, and other effects. Here, however, it is purely a space self-interaction effect. This new dynamical effect also explains the bore hole g anomalies, and the black hole "mass spectrum". Eqn.(4), even when $\rho = 0$, has an expanding universe Hubble solution that fits the recent supernovae data in a parameter-free manner without requiring "dark matter" nor "dark energy", and without the accelerating expansion artifact [4]. However (7) cannot be entirely expressed in terms of $\mathbf{g}$ because the fundamental dynamical variable is $\mathbf{v}$. The role of (7) is to reveal that if we analyse gravitational phenomena we will usually find that the matter density ρ is insufficient to account for the observed $\mathbf{g}$. Until recently this failure of Newtonian gravity has been explained away as being caused by some unknown and undetected "dark matter" density. Eqn.(7) shows that to the contrary it is a dynamical property of 3-space itself. Significantly the quantum matter 3-space-induced 'gravitational' acceleration in (6) also follows from maximising the elapsed proper time wrt the wave-packet trajectory $\mathbf{r}_o(t)$, see [2],

$$\tau=\int dt\sqrt{1-\frac{\mathbf{v}_R^2(\mathbf{r}_o(t),t)}{c^2}},\tag{9}$$

and then taking the limit $v_R/c\to 0$. This shows that (i) the matter 'gravitational' geodesic is a quantum wave refraction effect, with the trajectory determined by a Fermat maximised proper-time principle, and (ii) that quantum systems undergo a local time dilation effect. Significantly the time dilation effect in (9) involves matter motion wrt the dynamical 3-space, and not wrt the observer, and so distinguishing LR from SR. A full derivation of (9) requires the generalised Dirac equation, with the replacement $\partial/\partial t\to\partial/\partial t+\mathbf{v}\cdot\nabla$, as in (5). In differential form (9) becomes

$$d\tau^2=g_{\mu\nu}dx^\mu dx^\nu=dt^2-\frac{1}{c^2}(d\mathbf{r}-\mathbf{v}(\mathbf{r}(t),t)dt)^2,\tag{10}$$

which introduces a curved spacetime metric $g_{\mu\nu}$ that emerges from (4). However this spacetime has no ontological significance — it is merely a mathematical artifact, and as such hides the underlying dynamical 3-space. This induced metric is not determined by the Einstein-Hilbert equations, which originated as a generalisation of Newtonian gravity, but without the knowledge that a dynamical 3-space had indeed been detected by Michelson and Morley in 1887 by detecting light speed anisotropy. In special circumstances, and with $\alpha = 0$, they do yield the same effective spacetime metric. However the dynamics in (4) is more general, as noted above, and has passed more tests.

3 New Gravity Equation for a Spherically Symmetric System

For the case of zero vorticity the matter acceleration in (6) gives

$$\mathbf{g}(\mathbf{r}, t) = \frac{\partial \mathbf{v}}{\partial t} + \frac{\nabla v^2}{2} \tag{11}$$

For a time independent flow we introduce a generalised gravitational potential, which gives a microscopic explanation for that potential,

$$\Phi(\mathbf{r}) = -\frac{v^2}{2}. \tag{12}$$

For the case of a spherically symmetric and time independent inflow we set $\mathbf{v}(\mathbf{r}, t) = -\hat{\mathbf{r}}v(r)$, then (4) becomes, with $v' = dv/dr$,

$$\frac{\alpha}{2r}\left(\frac{v^2}{2r} + vv'\right) + \frac{2}{r}vv' + (v')^2 + vv'' = -4\pi G\rho \tag{13}$$

which can be written as

$$\frac{1}{r^2}\frac{d}{dr}\left(r^{2-\frac{\alpha}{2}}\frac{d}{dr}\left(r^{\frac{\alpha}{2}}\Phi\right)\right) = 4\pi G\rho \tag{14}$$

This form suggests that the new dynamics can be incorporated into the space metric, in that the 3-space α-term appears to lead to a fractal dimension of $3 - \alpha/2 = 2.996$, see [10]. The velocity flow description of space is completely equivalent to Newtonian gravity when the α dependent term in (4) is removed. In this case setting $\alpha = 0$ reduces (14) to the Poisson equation of Newtonian gravity for the case of spherical symmetry.

4 Solutions to New Gravity Equation for Non-Uniform Density

The solutions to (14) for a uniform density distribution are published in [2]. For variable density $\rho(r)$ the exact solution to (14) is*

$$\begin{aligned}\Phi(r) = - &\frac{\beta}{r^{\frac{\alpha}{2}}} - \frac{G}{(1-\frac{\alpha}{2})r}\int_0^r 4\pi s^2\rho(s)ds \\ - &\frac{G}{(1-\frac{\alpha}{2})r^{\frac{\alpha}{2}}}\int_r^\infty 4\pi s^{1+\frac{\alpha}{2}}\rho(s)ds,\end{aligned} \tag{15}$$

When $\rho(r) = 0$ for $r > R$, this becomes

$$\Phi(r) = \begin{cases} -\dfrac{\beta}{r^{\frac{\alpha}{2}}} - \dfrac{G}{(1-\frac{\alpha}{2})r}\displaystyle\int_0^r 4\pi s^2\rho(s)ds \\ -\dfrac{G}{(1-\frac{\alpha}{2})r^{\frac{\alpha}{2}}}\displaystyle\int_r^R 4\pi s^{1+\frac{\alpha}{2}}\rho(s)ds, \quad 0 < r \le R \\ -\dfrac{\beta}{r^{\frac{\alpha}{2}}} - \dfrac{\gamma}{r}, \qquad r > R \end{cases} \tag{16}$$

where

$$\gamma = \frac{G}{(1-\frac{\alpha}{2})}\int_0^R 4\pi s^2\rho(s)ds = \frac{GM}{(1-\frac{\alpha}{2})} \tag{17}$$

Here M is the total matter mass, and β is a free parameter. The term $\beta/r^{\alpha/2}$ describes an inflow singularity or "black hole" with arbitrary strength. This is unrelated to the putative black holes of General Relativity. This corresponds to a primordial black hole. As well the middle term in (16) also has a $1/r^{\alpha/2}$ inflow-singularity, but whose strength is mandated by the matter density, and is absent when $\rho(r) = 0$ everywhere. This is a minimal "black hole", and is present in all matter systems. The $\beta/r^{\alpha/2}$ term will produce a long range gravitational acceleration $g = \beta/r^{1+\alpha/2}$, as observed in spiral galaxies. For the region outside the sun ($r > R$) Keplerian orbits are known to well describe the motion of the planets within the solar system, apart from some small corrections, such as the Precession of the Perihelion of Mercury, which follow from relativistic effects from (9). Thus is the case $\beta = 0$, and the sun has only an induced 'Minimal Attractor'. These minimal black holes contribute to the external $g = K/r^2$ gravitational acceleration, through an effective mass

$$M_{BH} = \frac{M}{1-\frac{\alpha}{2}} - M = \frac{\alpha}{2}\frac{M}{1-\frac{\alpha}{2}} \approx \frac{\alpha}{2}M \tag{18}$$

as previously reported [2]. These induced black hole " effective" masses have been detected in numerous globular clusters and spherical galaxies and their predicted effective masses have been confirmed in some 19 such cases [11]. These gave the value $\alpha = \approx 1/137$ [12]. The induced black hole dynamics at the center of the sun is responsible for the new density, pressure and temperature profiles computed herein.

*Eqn (14) also permits a $-\overline{\gamma}/r$ term in (16). However this is not valid, as the full [3] version of (14) would then involve a point mass at $r = 0$, because $\nabla^2(1/|r|) = -4\pi\delta(\mathbf{r})$, and in (16) all the mass is accounted for by $\rho(r)$. See [2] for a detailed discussion.

5 Polytropic Models using Dynamical 3-Space Theory

For a star to be in hydrostatic equilibrium the inward force of gravity must match the net outward effect of the pressure,

$$\frac{dP}{dr} = -\frac{d\Phi}{dr}\rho \tag{19}$$

Here we use the polytrope modelling of the pressure-density equation of state.

$$P = K\rho^{1+\frac{1}{n}} \tag{20}$$

where n is the polytropic index, and K is a constant. This was introduced by Eddington, and was extensively used by Chandrasekhar [13–16], but these analyses only apply in the case of Newtonian gravity. The new theory of gravity requires a new treatment.

The polytropic relation between pressure and density (20) gives

$$\frac{dP}{dr} = \frac{K(n+1)}{n}\rho^{\frac{1}{n}}\frac{d\rho}{dr} \tag{21}$$

and (19) gives

$$\frac{d\Phi}{dr} = -\frac{K(n+1)}{n}\rho^{\frac{1}{n}-1}\frac{d\rho}{dr} \tag{22}$$

Integration gives

$$\Phi = -K(n+1)\rho^{\frac{1}{n}} + C \tag{23}$$

Here it will be useful to define the gravitational potential at the sun's surface $\Phi_R = \Phi(R) = C$ as the value of the integration constant, and so we obtain for the density

$$\rho = \left(\frac{\Phi_R - \Phi}{K(n+1)}\right)^n \tag{24}$$

One of the characteristics of the new gravity is that all spherical objects contain induced black holes. In the context of polytrope models this presents the problem that the central value of the potential cannot be used, as in the Lane-Emden equation. We can however impose the polytropic condition from (24) onto numerical solutions to iteratively solve the problem. Multiplying (24) by $4\pi r^2$ and integrating yields

$$M = \int_0^R 4\pi r^2 \rho dr = \int_0^R 4\pi r^2 \left(\frac{\Phi_R - \Phi)}{K(n+1)}\right)^n dr \tag{25}$$

and then

$$K = \frac{1}{(n+1)M^{1/n}}\left(\int_0^R 4\pi r^2(\Phi_R - \Phi)^n dr\right)^{1/n} \tag{26}$$

A new density distribution and K value can now be calculated from an initial density distribution by cycling through the following relations iteratively

$$\begin{aligned}\Phi(r) =& \frac{-G}{(1-\frac{\alpha}{2})}\left(\frac{1}{r}\int_0^r 4\pi s^2\rho(s)ds + \right.\\ &\left. + \frac{1}{r^{\frac{\alpha}{2}}}\int_r^R 4\pi s^{1+\frac{\alpha}{2}}\rho(s)ds\right)\\ K =& \frac{1}{(n+1)M^{1/n}}\left(\int_0^R 4\pi r^2(\Phi_R - \Phi)^n dr\right)^{1/n}\\ \rho(r) =& \left(\frac{\Phi_R - \Phi(r)}{K(n+1)}\right)^n\end{aligned} \tag{27}$$

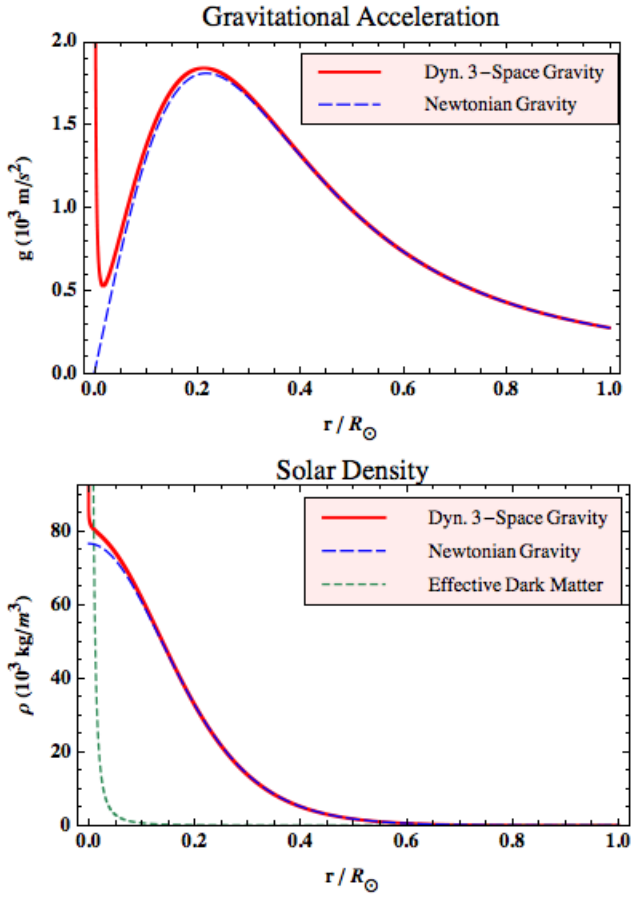


Fig. 1: Gravity and density plots for a polytropic model for the sun with $n = 3$. The effective dark matter distribution is shown in the density plot.

6 Polytropic Solar Models

For the sun a polytrope model with $n = 3$ is known to give a good approximation to conditions in the solar core as compared with the Standard Solar Model [16]. This is known as the Eddington Standard Model. The polytrope model does well in comparison with the Standard Solar Model [17]. To test the calculation method, setting $\alpha = 0$ should reproduce

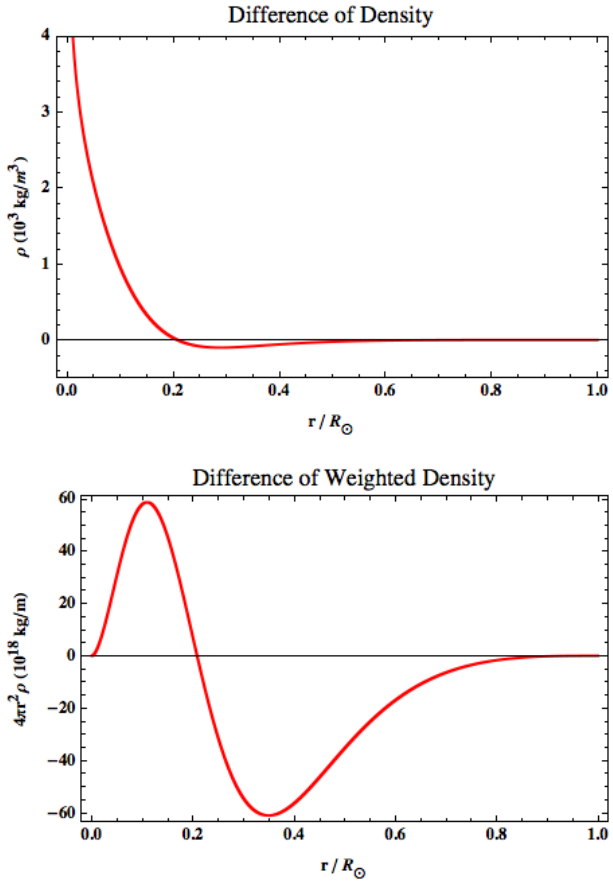


Fig. 2: Top graph shows difference in density $\rho(r)$ between new gravity and Newtonian modeling. Bottom graph show the difference in weighted density $4\pi r^2\rho(r)$.

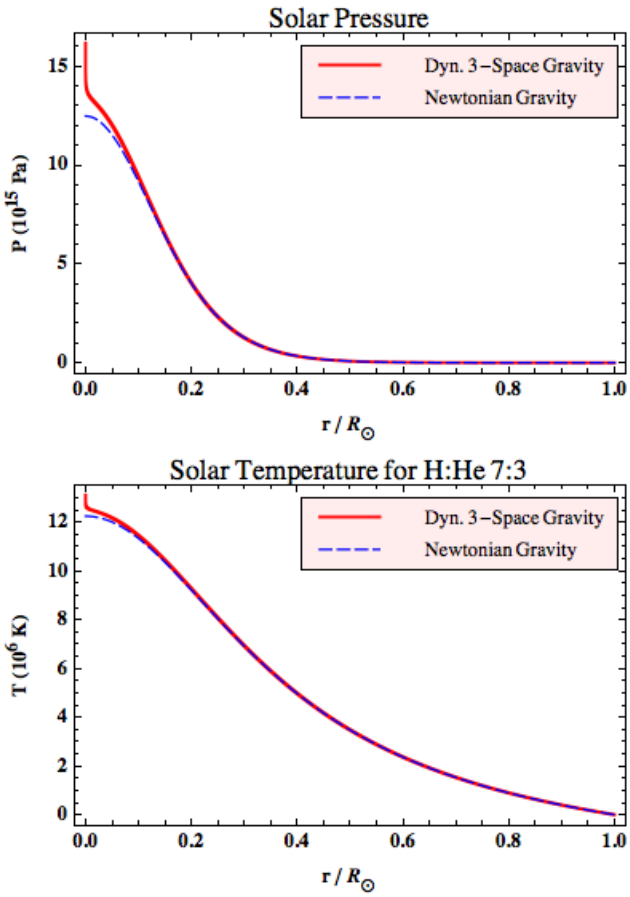


Fig. 3: The pressure and temperature in the center of the sun is predicted to be much larger in the new model.

the results from the Lane-Emden equation, which is based on Newtonian gravity. The results of starting with a uniform density and then iteratively finding the solution agree with the values published by Chandrasekhar [13]. The density distribution also matched numerical solutions produced in Mathematica to the Lane-Emden equation.

Results from solving the equations in (27) iteratively, until convergence was achieved, are shown in Figs.1-3 for various quantities, and compared with the results for Newtonian gravity. For the new gravity ($\alpha = 1/137$) we see a marked increase in the gravity strength $g(r)$ near the center, Fig.1, caused by the induced black hole at the center, which is characteristic of the new gravity theory, and which draws in the matter to enhance the matter density near the center.

The new model of gravity has been used to explain away the need for dark matter in astrophysics [4]. Here we find the effective "dark matter" distribution that would need to be added to the matter distribution to create these gravitational effects in Newtonian Gravity. From (8) and (12) we obtain

$$\rho_{DM}(r) = -\frac{\alpha}{8\pi G r}\left(\frac{\Phi}{r} + \frac{d\Phi}{dr}\right). \qquad (28)$$

Using (16) we then obtain

$$\rho_{DM}(r) = \frac{\alpha}{2} r^{-2-\alpha/2} \int_r^R s^{1+\alpha/2}\rho(s)ds. \qquad (29)$$

This effective "dark matter" distribution is shown in Fig.1 for the polytropic sun model. This then gives the total "dark matter"

$$M_{DM} = \int_0^R 4\pi r^2 \rho_{DM}(r)dr = \frac{\alpha}{2}\frac{M}{1-\frac{\alpha}{2}} \qquad (30)$$

in agreement with (18). The "dark matter" effect is the same as the induced "black hole" effect, in the new gravity theory.

The matter density has increased towards the center, as seen in Fig.2, and so necessarily there is a slightly lower matter density in the inner middle region. This effect is more clearly seen in the plot of $4\pi r^2\rho(r)$. The "dark matter"/"black hole" effect contributes to the external gravitational acceleration, and so the total mass of the sun, defined as its matter content, is lower than computed using Newtonian gravity, see (17). The total mass is now 0.37% ($\equiv \alpha/(2-\alpha)$) smaller.

The pressure and temperature generated by the new gravity is shown in Fig.3. The pressure comes from the poly-

trope relation, (20), and closely follows the density distribution. The temperature can be calculated from the ideal gas equation, with $\mu = 0.62$ corresponding to a ratio of 7:3 of Hydrogen to Helium, to obtain

$$T(r) = \frac{Pm_p\mu}{k\rho} \tag{31}$$

where m_p is the mass of a proton, k is Boltzmann's constant and μ is the mass ratio. Unlike the pressure and density, the temperature is increased in the middle region as well as the inner region.

7 Conclusions

The discovery of the dynamical 3-space changes most of physics. This space has been repeatedly detected in light-speed anisotropy experiments. The dynamics of this space follow from a unique generalisation of Newtonian gravity, once that is expressed in a velocity framework. Then the gravitational acceleration field $\mathbf{g}(\mathbf{r}, t)$ is explained as the local acceleration of the structured space, with evidence that the structure is fractal. This space is the local absolute frame of reference. Uniquely incorporating this space into a generalised Schrödinger equation shows that, up to vorticity effects and relativistic effects, the quantum matter waves are refracted by the space, and yield that quantum matter has the same acceleration as that of space itself. So this new physics provides a quantum theory derivation of the phenomenon of gravity. The 3-space dynamics involves G and the fine structure constant α, with this identification emerging from the bore hole gravity anomalies, and from the masses of the minimal "black holes" reported for globular clusters and spherical galaxies. There are numerous other phenomena that are now accounted for, including a parameter-free account of the supernova red-shift — magnitude data. The occurrence of α implies that we are seeing evidence of a new unified physics, where space and matter emerge from a deeper theory. One suggestion for this theory is *Process Physics*.

Herein we have reported the consequences of the new, emergent, theory of gravity, when applied to the sun. This theory predicts that the solar core, which extends to approximately 0.24 of the radius, is hotter, more dense and of higher pressure than current Newtonian-gravity based models. Thus a new study is now needed on how these changes will affect the solar neutrino output. It is also necessary to revisit the stellar evolution results.

Submitted on August 23, 2010 / Accepted on September 20, 2010

References

1. Newton I. Philosophiae Naturalis Principia Mathematica, 1687.
2. Cahill R. T. Process Physics: From Information Theory to Quantum Space and Matter. Nova Science Pub., New York, 2005.
3. Cahill R. T. Dynamical 3-Space: A Review, in: Ether space-time and cosmology: New insights into a key physical medium, Duffy M. and Lévy J. (Editors), Apeiron, 2009, pp. 135-200.
4. Cahill R. T. Unravelling the Dark Matter — Dark Energy Paradigm, *Apeiron*, 2009, v. 16, no. 3, 323–375.
5. Cahill R. T., Kitto K. Michelson-Morley Experiments Revisited. *Apeiron*, 2003, v. 10, no. 2, 104–117.
6. Cahill R. T. The Michelson and Morley 1887 Experiment and the Discovery of Absolute Motion. *Progress in Physics*, 2005, v. 3, 25–29.
7. Cahill R. T. Combining NASA/JPL One-Way Optical-Fiber Light-Speed Data with Spacecraft Earth-Flyby Doppler-Shift Data to Characterise 3-Space Flow. *Progress in Physics*, 2009, v. 4, 50–64.
8. Cahill R. T. Unravelling Lorentz Covariance and the Spacetime Formalism.*Progress in Physics*, 2008, v. 4, 19–24.
9. Bahcall J. N. Neutrino Astrophysics. Cambridge University Press, Cambridge, England, 1989.
10. Cahill R. T. Dynamical Fractal 3-Space and the Generalised Schrödinger Equation: Equivalence Principle and Vorticity Effects. *Progress in Physics*, 2006, v. 1, 27–34.
11. Cahill R. T. Black Holes and Quantum Theory: The Fine Structure Constant Connection. *Progress in Physics*, 2006, v. 4, 44–50.
12. Cahill R. T. 3-Space Inflow Theory of Gravity: Boreholes, Blackholes and the Fine Structure Constant. *Progress in Physics*, 2006, v. 2, 9–16.
13. Chandrasekhar S. An Introduction to the Study of Stellar Structure. Dover Publications, New York, 1958.
14. Hansen C. J., Kawaler S. D. Trimble V. Stellar Interiors — Physical Principles, Structure and Evolution. Springer, New York, 2004.
15. Horedt G. P. Polytropes. Applications in Astrophysics and Related Fields. Kluwer, Dordrecht, 2004.
16. Kippenhahn R., Weigert A. Stellar Structure and Evolution. Springer, New York, 1994.
17. Bahcall J. N., Basu S., Pinsonneault M. H. How Uncertain are Solar Neutrino Predictions? *Physics Letters B*, 1998, v. 433, 1–8.

Particles and Antiparticles in the Planck Vacuum Theory

William C. Daywitt

National Institute for Standards and Technology (retired), Boulder, Colorado, USA, E-mail: wcdaywitt@earthlink.net

This short note sheds some light on the negative energy vacuum state by expanding the Planck vacuum (PV) model and taking a closer look at the particle-antiparticle nature of the Dirac equation. Results of the development are briefly discussed with regard to the complexity of the PV interaction with the massless free charge, the Dirac electron, and the proton; an exercise that may lead to a better proton model.

The negative energy PV model [1] can be expanded to include negative energy particle states in the following manner: the structure of the PV is related to the string of Compton relations

$$r_e m_e c^2 = \cdots = r_p m_p c^2 = \cdots = r_* m_* c^2 = e_*^2 = c\hbar \qquad (1)$$

where the subscripts represent respectively the electron, proton, Planck particle, and their antiparticles; and where the dots represent any number of intermediate particle-antiparticle states. The r_e and m_e, etc., are the Compton radii and masses of the various particles, c is the speed of light, and $\hbar$ is Planck's constant. The bare charge e_* is assumed to be massless and is related to the elementary charge e observed in the laboratory via $e^2 = \alpha e_*^2$, where α is he fine structure constant. The particle-antiparticle masses are the result of their bare charges being driven by ultra-high-frequency zero-point fields that exist in free space [2, 3]. The charge on the Planck particles within the PV is negative. It is assumed that positive charges are holes that exist within the negative energy PV, an assumption that is supported by the Dirac equation and its negative energy solution [4].

The relation of positive and negative particles and antiparticles to the Compton relations in (1) is easily explained. In the above scheme, negatively charged particles or antiparticles exist in free space and exert a perturbing force [1]

$$\frac{(-e_*)(-e_*)}{r^2} - \frac{mc^2}{r} \qquad (2)$$

on the PV, where m is the particle-antiparticle mass. The first charge on the left is due to the free particle or antiparticle and the second to the Planck particles within the PV. The hole exerts a corresponding force within the PV equal to

$$\frac{(+e_*)(-e_*)}{r^2} - \frac{(-mc^2)}{r} \qquad (3)$$

where the effective positive charge on the left is due to the missing negative charge (the hole) in the PV sea and the negative mass energy $(-mc^2)$ is due to the hole belonging to a negative energy state. The radius r at which (2) and (3) vanish is the particle or antiparticle Compton radius $r_c\,(= e_*^2/mc^2)$. The more complete form for (1) can then be expressed as

$$r_e(\pm\, m_e c^2) = \cdots = r_p(\pm\, m_p c^2) = \cdots = r_*(\pm\, m_* c^2) = \pm\; e_*^2 \qquad (4)$$

which renders its application to both particles and antiparticles more explicit and transparent. The positive mass energies belong to the negatively charged free-space particles or antiparticles, while the negative mass energies belong to the PV holes which are responsible for the ficticious positively charged particles or antiparticles imagined to exist in free space. Both equations in (4) lead back to the single equation (1) which defines $\hbar$.

The preceding ideas are illustrated using the Dirac equation and provide a clearer view of that equation as it is related to the concept of Dirac holes. The Dirac equation for the electron can be expressed as [4, 5]

$$\left(c\,\vec{\alpha}\cdot\widehat{p_e} + \beta m_e c^2\right)\psi_e = E_e\psi_e \qquad (5)$$

where the momentum operator and energy are given by

$$\widehat{p_e} = \frac{\hbar\nabla}{i} = \frac{(-e_*)(-e_*)\nabla}{ic} \text{ and } E_e = +\sqrt{m_e^2c^4 + c^2p_e^2} \qquad (6)$$

and where $\vec{\alpha}$ and β are defined in [5]. The relativistic momentum is $p_e\,(= m_e v/\sqrt{1 - v^2/c^2}\,)$. The shift from the positive-energy electron solution to the negative-energy hole (positron) solution proceeds as follows:

$$E_e \longrightarrow E_h = -E_e \qquad (7)$$

$$m_e c^2 \longrightarrow m_h c^2 = -m_e c^2 \qquad (8)$$

$$p_e = \frac{m_e v}{\sqrt{1 - v^2/c^2}} \rightarrow p_h = \frac{-m_h v}{\sqrt{1 - v^2/c^2}} = -p_e, \qquad (9)$$

$$\widehat{p_e} = \frac{(-e_*)(-e_*)\nabla}{ic} \rightarrow \widehat{p_h} = \frac{(+e_*)(-e_*)\nabla}{ic} = -\widehat{p_e}. \qquad (10)$$

Substituting equations (7) through (10) into (5) yields

$$\left(c\,\vec{\alpha}\cdot\widehat{p_h} + \beta m_h c^2\right)\psi_h = E_h\psi_h \qquad (11)$$

for the hole solution, where $E_h = -(m_h^2c^4 + c^2p_h^2)^{1/2}$. From (5) and (11) and $m_h = m_e$ it follows that the electron and hole satisfy the same Dirac equation of motion with $E_h = -E_e$. Although the hole exists in the PV, it appears experimentally in free space as a positron due to the hole's field permeating that space. In turn, the positron's deflection in a free-space magnetic field is due to that field permeating the PV and affecting the hole.

From the development of the PV theory so far, the Dirac equation appears to be part of a succession of equations involving an increasingly more complicated interaction between the free-space particle and the PV. For example, the interaction of a massless point charge traveling at a constant velocity results in the relativistic electric and magnetic fields (and by inference the Lorentz transformation) that can be easily calculated directly from the charge's Coulomb field (the first term in (2)) and its interaction with the PV [1, Section 4]. The Dirac electron (a massive point charge) is next in complexity to the point charge and perturbs the PV with the total force in (2), leading to the Dirac equation (and the quantum fields associated with it) which represents the PV reaction to the moving Dirac electron [4].

The proton is the next more complex and stable particle whose properties are shaped by its interaction with the PV. Being in essence a more complicated PV hole than the positron, the proton exhibits some structure as witnessed by its three-quark nature associated (it seems correct to assume) with the hole. The calculational difficulties besetting quantum chromodynamics [6, p.70] attest to the idea expressed above that things are getting more complex in the progression from leptons to hadrons and their PV interactions. Perhaps these difficulties can be resolved by a better model for the heavy particles based on the PV theory.

Submitted on September 5, 2010 /Accepted on September 7, 2010

References

1. Daywitt W. C. The Planck vacuum. *Progress in Physics*, 2009, v. 1, 20–26.
2. Puthoff H. E. Gravity as a zero-point-fluctuation force. *Physical Review A*, 1989, v. 39, no. 5, 2333–2342.
3. Daywitt W. C. The source of the quantum vacuum. *Progress in Physics*, 2009, v. 1, 27–32.
4. Daywitt W. C. The Dirac electron in the Planck vacuum theory. *Progress in Physics*, 2010, v. 4, 69–71.
5. Gingrich D. M. Practical quantum electrodynamics. CRC, The Taylor & Francis Group, Boca Raton, 2006.
6. Giunti C., Kim C. W. Fundamentals of neutrino physics and astrophysics. Oxford Univ. Press, Oxford, 2007.

Wave Particle Duality and the Afshar Experiment

Aurélien Drezet
Institut Neel, 25 rue des Martyrs 38042, Grenoble, France. E-mail: aurelien.drezet@grenoble.cnrs.fr

We analyze the experiment realized in 2003-2004 by S. Afshar et al. [1] in order to refute the principle of complementarity. We discuss the general meaning of this principle and show that contrarily to the claim of the authors Bohr's complementarity is not in danger in this experiment.

1 Introduction

In an interesting series of articles published few years ago Afshar and coworkers [1,2] reported an optical experiment in which they claimed to refute the well known N. Bohr principle of complementarity [3,4]. Obviously this result, if justified, would constitute a serious attack against the orthodox interpretation of quantum mechanics (known as the Copenhagen interpretation). This work stirred much debate in different journals (see for examples references [5–12]).

We think however that there are still some important misunderstandings concerning the interpretation of this experiment. In a preprint written originally in 2004 [5] (and following some early discussions with Afshar) we claimed already that the interpretation by Afshar *et al.* can be easily stated if we stay as close as possible from the texts written by Bohr. The aim of the present article (which was initially written in 2005 to precise a bit the thought developed in [5]) is to comment the interpretation discussed in [1]. We will in the following analyze the meaning of Bohr principle and show that far from disproving its content the experiment [1] is actually a complete confirmation of its general validity.

The difficulties associated with the understanding of this principle are not new and actually complementarity created troubles even in Einstein mind [3] so that we are here in good company. To summarize a bit emphatically Bohr's complementarity we here remind that this principle states that if one of a pair of non commuting observables of a quantum object is known for sure, then information about the second (complementary) is lost [3, 4, 15, 16]. This can be equivalently expressed as a kind of duality between different descriptions of the quantum system associated with different experimental arrangements which mutually exclude each other (read in particular [3,4]). Later in the discussion we will try to precise this definition but for the moment it is enough to illustrate the concepts by examples

Consider for instance the well known Young double-pinholes interference experiment made with photons. The discrete nature of light precludes the simultaneous observation of a same photon in the aperture plane and in the interference pattern: the photon cannot be absorbed twice. This is already a trivial manifestation of the principle of Bohr. Here it implies that the two statistical patterns associated with the wave in the aperture plane and its Fourier (i. e., momentum) transform require necessarily different photons for their recording. It is in that sense that each experiment excludes and completes reciprocally the other. In the case considered before the photon is absorbed during the first detection (this clearly precludes any other detection). However even a non-destructive solution for detection implying entanglement with other quantum systems has a radical effect of the same nature: the complementarity principle is still valid. For example, during their debate Bohr and Einstein [3] discussed an ideal *which-way* experiment in which the recoil of the slits is correlated to the motion of the photon. Momentum conservation added to arguments based on the uncertainty relations are sufficient to explain how such entanglement photon-slits can erase fringes [15–19]. It is also important for the present discussion to remind that the principle of complementarity has a perfidious consequence on the experimental meaning of trajectory and path followed by a particle. Indeed the unavoidable interactions existing between photons and detectors imply that a trajectory existing independently of any measurement process cannot be unambiguously defined. This sounds even like a tragedy when we consider once again the two-holes experiment. Indeed for Bohr this kind of experiments shows definitely the essential element of ambiguity which is involved in ascribing conventional physical attributes to quantum systems. Intuitively (i. e., from the point of view of classical particle dynamic) one would expect that a photon detected in the focal plane of the lens must have crossed only one of the hole 1 or 2 before to reach its final destination. However, if this is true, one can not intuitively understand how the presence of the second hole (through which the photon evidently did not go) forces the photon to participate to an interference pattern (which obviously needs an influence coming from both holes). Explanations to solve this paradox have been proposed by de Broglie, Bohm, and others using concepts such as empty waves or quantum potentials [20, 21]. However all these explanations are in agreement with Bohr principle (since they fully reproduce quantum predictions) and can not be experimentally distinguished. Bohr and Heisenberg proposed for all needed purposes a much more pragmatic and simpler answer: *don't bother*, the complementarity principle precludes the simultaneous observation of a photon trajectory and of an interference pattern. For Bohr [3]: *This point is of great logical consequence, since it is only the circumstance that we are presented with a choice of either tracing the path*

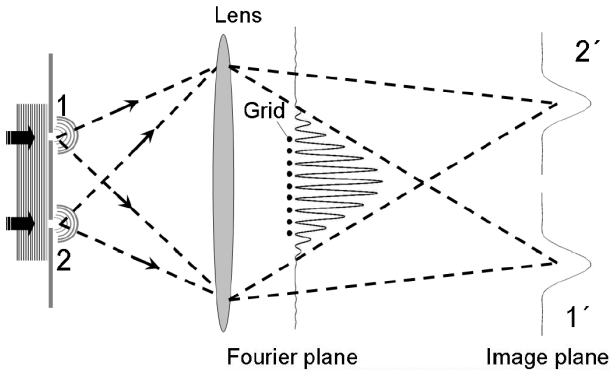


Fig. 1: The experiment described in [1]. Photons coming from pinholes 1 and 2 interfere in the back-focal plane of a lens (Fourier plane) whereas they lead to two isolated narrow spots in the image plane (the image plane is such that its distance p' to the lens is related to the distance p between the lens and the apertures screen by $1/p + 1/p' = 1/f$, where f is the focal length). The wire grid in the back focal plane, distant of f from the lens, is passing through the minima of the interference pattern. The subsequent propagation of the wave is consequently not disturbed by the grid.

of a particle or observing interference effects, which allows us to escape from the paradoxical necessity of concluding that the behaviour of an electron or a photon should depend on the presence of a slit in the diaphragm through which it could be proved not to pass. From such an analysis it seems definitively that Nature resists to deeper experimental investigation of its ontological level. As summarized elegantly by Brian Greene [22]: *Like a Spalding Gray soliloquy, an experimenter's bare-bones measurement are the whole show. There isn't anything else. According to Bohr, there is no backstage*. In spite of its interest it is however not the aim of the present article to debate on the full implications of such strong philosophical position.

2 Complementarity versus the experiments

2.1 A short description of the Afshar *et al.* experiment

The experiment reported in [1] (see Fig. 1) is actually based on a modification of a *gedanken* experiment proposed originally by Wheeler [23]. In the first part of their work, Afshar *et al.* used an optical lens to image the two pinholes considered in the Young interference experiment above mentioned. Depending of the observation plane in this microscope we can then obtain different complementary information.

If we detect the photons in the focal plane of the lens (or equivalently just in front of the lens [24]) we will observe, i.e, after a statistical accumulation of photon detection events, the interference fringes. However, if we record the particles in the image plane of the lens we will observe (with a sufficiently high numerical aperture) two sharp spots 1' and 2' images of the pinholes 1 and 2. Like the initial Young two-holes experiment this example illustrates again very well the principle of Bohr. One has indeed complete freedom for measuring the photon distribution in the image plane instead of detecting the fringes in the back focal plane. However, the two kinds of measurements are mutually exclusive: a single photon can participate only to one of these statistical patterns.

In the second and final part of the experiment, Afshar *et al.* included a grid of thin absorbing wires located in the interference fringes plane. Importantly, in the experiment the wires must be located at the minimum of the interference pattern in order to reduce the interaction with light. In the following we will consider a perfect interference profile (with ideal unit visibility $V = (I_{max} - I_{min})/(I_{max} + I_{min}) = 1$) to simplify the discussion. If additionally the geometrical cross section of each wire tends ideally to zero then the interference behavior will, at the limit, not be disturbed and the subsequent wave propagation will be kept unchanged. This implies that the photon distributions 1' and 2', located in the image plane optically conjugated with the aperture plane, are not modified by the presence, or the absence, of the infinitely thin wire grid. Naturally, from practical considerations an infinitely thin dielectric wire is not interacting with light and consequently produces the same (null) effect whatever its location in the light path (minimum or maximum of the interference for example). In order to provide a sensible probe for the interference pattern, necessary for the aim of the experiment considered, we will suppose in the following idealized wires which conserve a finite absorbtion efficiency and this despite the absence of any geometrical transversal extension. We will briefly discuss later what happens with spatially extended scattering wires with finite cross section, but this point is not essential to understand the essential of the argumentation. With such wires, and if we close one aperture (which implies that there is no interference fringes and thus that a finite field impinges on the wires) the scattering and absorbtion strongly affect the detection behavior in the image plane. As it is seen experimentally [1, 2] the scattering by the wire grid in general produces a complicated diffraction pattern and not only an isolated narrow peak in 1' or 2' as it would be without the grid.

In such conditions, the absence of absorbtion by the wires when the two apertures are open is a clear indication of the existence of the interference fringes zeros, i.e., of a wave-like character, and this even if the photon is absorbed in the image plane in 1' or 2'. Following Afshar *et al.*, this should be considered as a violation of complementarity since the same photons have been used for recording *both* the 'path' and the wave-like information. The essential questions are however what we mean precisely here by path and wave-like information and what are the connections of this with the definition of complementarity. As we will see hereafter it is by finding a clear answer to these questions that the paradox and the contradictions with Bohr's complementarity are going to vanish.

2.2 The wave-particle duality mathematical relation

At that stage, it is important to point out that the principle of complementarity is actually a direct consequence of the mathematical formalism of quantum mechanics and of its statistical interpretation [4]. It is in particular the reason why the different attempts done by Einstein to refute complementarity and the Heisenberg uncertainty relations always failed: the misinterpretations resulted indeed from a non-cautious introduction of classical physics in the fully consistent quantum mechanic formalism [3]. For similar reasons here we show that a problem since Afshar *et al.* actually mixed together, i.e imprudently, argumentations coming from classical and quantum physics. We will show that this mixing results into an apparent refutation of the complementarity principle.

After this remark we now remind that a simple mathematical formulation of complementarity exists in the context of two path interferometry [25–28]. For example in the Young double-apertures experiment considered previously the field amplitudes C_1 and C_2 associated with the two narrow apertures, separated by the distance d, allow us to define the wave function in the two-apertures plane by:

$$\psi(x) \sim C_1\delta(x - d/2) + C_2\delta(x + d/2). \quad (1)$$

From this formula one can easily introduce the "distinguishability"

$$K = \frac{||C_1|^2 - |C_2|^2|}{|C_1|^2 + |C_2|^2}. \quad (2)$$

This quantity can be physically defined by recording the photons distribution in the aperture plane and constitutes an observable measure of the "path" distinguishability (see however section 3.3). The interpretation of K is actually clear, and in particular if $K = 0$ each apertures play a symmetrical role, whereas if $K = 1$ one of the two apertures is necessarily closed. Naturally, like in the Afshar experiment, K can also be measured by recording photons in the image plane of the lens in 1' and 2'. Equations (1) and (2) are still valid, with the only differences that: i) we have now a diffraction spot (like an Airy disk) instead of a Dirac distribution in equation (1), and ii) that the spatial variables are now magnified by the lens.

Instead of the spatial representation one can also consider the Fourier transform corresponding to the far field interference pattern recorded at large distance of the two-slits screen:

$$\psi(k) \sim C_1 \cdot e^{ikd/2} + C_2 \cdot e^{-ikd/2}. \quad (3)$$

Such a wave is associated with an oscillating intensity in the k-space given by

$$I(k) \sim 1 + V\cos(kd + \chi) \quad (4)$$

where $\chi = \arg(C_1) - \arg(C_2)$ and V is the fringe visibility

$$V = \frac{2|C_1| \cdot |C_2|}{|C_1|^2 + |C_2|^2}. \quad (5)$$

This quantity is also a physical observable which can defined by recording the photons in the far-field, or, like in the Afshar *et al.* first experiment, by recording the photons fringes in the back focal plane of the lens (the back focal plane is the plane where the momentum distribution $\hbar k$ is experimentally and rigorously defined [16]). Like it is for K, the meaning of V is also very clear: if $V = 1$ both apertures must play a symmetrical role, whereas if $V = 0$ only one aperture is open.

A direct mathematical consequence of equations (2) and (5) is the relation

$$V^2 + K^2 = 1, \quad (6)$$

which expresses the duality [25, 26] between the two mathematical measures K and V associated with the two mutually exclusive (i.e., complementary) experiments in the direct and Fourier space respectively. A particularly important application of equation (6) concerns which-path experiments. In such experiments, we wish to observe the interference pattern, and to find through each hole each photon is going through. As we explained before, a photon can not be observed twice, and this represents in general a fatal end for such expectations. There is however an important exception in the particular case with only one aperture open (i.e., $K = 1$). Indeed, in such case it is not necessary to record the photon in the aperture plane to know its path since if it is detected (in the back focal plane) it necessarily means that it went through the opened aperture. Of course, from equation (6) we have in counterpart $V = 0$, which means that fringes are not possible.

This dilemma, can not be solved by considering less invasive methods, like those using entanglement between the photon and an other quantum system or an internal degree of freedom (such as polarization or spins). To see that we consider a wave function $|\Psi\rangle$ describing the entanglement between the photon and these others quantum variables defining a which-path detector. We write

$$|\Psi\rangle = \int [C_1\delta(x - d/2)|x\rangle|\gamma_1\rangle + C_2\delta(x + d/2)|x\rangle|\gamma_2\rangle]dx$$
$$= \int [C_1 \cdot e^{ikd/2}|k\rangle|\gamma_1\rangle + C_2 \cdot e^{-ikd/2}|k\rangle|\gamma_2\rangle]dp \quad (7)$$

where $|\gamma_1\rangle$ and $|\gamma_2\rangle$ are the quantum state of the which-path detector if the photon is going through the aperture 1 or 2. Consider now the kind of information one can extract from $|\Psi\rangle$. First, by averaging (tracing) over the detector degrees of freedom we can define the total probability $P(x) = \mathrm{Tr}[\hat{\rho}|x\rangle\langle x|]$ of detecting a photon in the aperture plane in x by

$$P(x) \propto |C_1|^2\langle\gamma_1|\gamma_1\rangle(\delta(x - d/2))^2$$
$$+|C_2|^2\langle\gamma_2|\gamma_2\rangle(\delta(x + d/2))^2. \quad (8)$$

with $\hat{\rho} = |\Psi\rangle\langle\Psi|$ is the total density matrix. By analogy with equation (2) the total distinguishability is then defined by

$$K = \frac{||C_1|^2\langle\gamma_1|\gamma_1\rangle - |C_2|^2\langle\gamma_2|\gamma_2\rangle|}{|C_1|^2\langle\gamma_1|\gamma_1\rangle + |C_2|^2\langle\gamma_2|\gamma_2\rangle}. \quad (9)$$

Same as for equations (3-5) we can define the total probability to detect a photon of (transverse) wave vector k by

$$P(k) = \mathrm{Tr}[\hat{\rho}|k\rangle\langle k|] \propto 1 + V\cos(kx + \phi), \tag{10}$$

where the visibility V is written

$$V = \frac{2|C_1| \cdot |C_2| \cdot |\langle\gamma_1|\gamma_2\rangle|}{|C_1|^2\langle\gamma_1|\gamma_1\rangle + |C_2|^2\langle\gamma_2|\gamma_2\rangle}. \tag{11}$$

By combining V and K we deduce immediately $K^2 + V^2 = \eta^2 \leq 1$ with

$$\eta^2 = 1 - \frac{4|C_1|^2 \cdot |C_2|^2 \cdot (\langle\gamma_1|\gamma_1\rangle\langle\gamma_2|\gamma_2\rangle - |\langle\gamma_1|\gamma_2\rangle|^2)}{(|C_1|^2\langle\gamma_1|\gamma_1\rangle + |C_2|^2\langle\gamma_2|\gamma_2\rangle)^2} \tag{12}$$

and where the inequality results from the Cauchy-Schwartz relation $\langle\gamma_1|\gamma_1\rangle\langle\gamma_2|\gamma_2\rangle - |\langle\gamma_1|\gamma_2\rangle|^2 \geq 0$.
However, we can remark that by tracing over the degrees of freedom associated with the detector we did not consider a which-path experiment but simply decoherence due to entanglement. In order to actually realize such a which-path experiment we need to calculate the joint probability associated with a recording of the photon in the state $|x\rangle$ (or $|k\rangle$) in coincidence with a measurement of the detector in the eigenstate $|\lambda\rangle$ corresponding to one of its observable. These joint probabilities read $P(x,\lambda) = \mathrm{Tr}[\hat{\rho}|x\rangle\langle x||\lambda\rangle\langle\lambda|]$ and $P(k,\lambda) = \mathrm{Tr}[\hat{\rho}|k\rangle\langle k||\lambda\rangle\langle\lambda|$ with

$$\begin{aligned} P(x,\lambda) &\propto |C_1|^2|\langle\lambda|\gamma_1\rangle|^2(\delta(x - d/2))^2 \\ &+|C_2|^2|\langle\lambda|\gamma_2\rangle|^2(\delta(x + d/2))^2 \\ P(k,\lambda) &=\propto 1 + V_\lambda\cos(kx + \phi_\lambda). \end{aligned} \tag{13}$$

Indeed, the aim of such entanglement with a degree of freedom $|\lambda\rangle$ (produced for example by inserting polarization converters like quarter or half wave-plates just after the apertures [32]) is to generate a wave function

$$\psi_\lambda(x) \sim C_{1,\lambda}\delta(x - d/2) + C_{2,\lambda}\delta(x + d/2) \tag{14}$$

with either $C_{1,\lambda}$ or $C_{2,\lambda}$ (but not both) equal to zero. A subsequent projection on $|\lambda\rangle$ will reveal the path information. However, from the duality relation given by equation (5) applied to $\psi_\lambda(x)$ it is now obvious that we did not escape from the previous conclusion. Indeed, while the photon was not destroyed by the entanglement with the which-path detector, we unfortunately only obtained path distinguishability ($K_\lambda = 1$) at the expense of losing the interference behavior ($V_\lambda = 0$).
From all these experiments, it is clear that the discreteness of photon, and more generally of every quantum object, is the key element to understand complementarity. This was evident without entanglement, since the only way to observe a particle is to destroy it. However, even the introduction of a 'which-path' quantum state $|\lambda\rangle$ does not change the rule of the game, since at the end of journey we necessarily need to project, that is to kill macroscopically, the quantum system. This fundamental fact, was already pointed out many times by Bohr in his writings when he considered the importance of separating the macroscopic world of the observer from the microscopic quantum system observed, and also when he insisted on the irreversible act induced by the observer on the quantum system during any measurement process [4].

Let now return to the interpretation of Afshar *et al*. experiments. In the configuration with the lens and without the grid, we have apparently a new aspect of the problem since the fringes occur in a plane located before the imaging plane. Contrarily to the which-path experiments above mentioned, where the destructive measurements occurred in the interference plane, we have a priori here the freedom to realize a 'fringes-interaction free-experiment' which aim is to observe the fringes without detecting the particle in the back focal plane whereas the destructive measurement will occur in the image plane (i.e., in 1' or 2'). The role of the grid is expected to provide such information necessary for the interference reconstruction. Due to the absence of disturbance by the grid, Afshar *et al*. logically deduce that the field equals zero at the wires locations. If we *infer* the existence of an interference pattern with visibility V we must have

$$V = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})} = \frac{(I_{max} - 0)}{(I_{max} + 0)} = 1, \tag{15}$$

since $I_{min} = 0$. This means that we can obtain the value of the visibility only from the two assumptions that (i) the form of the profile should be a 'cos' function given by equation (4), and that (ii) no photon have been absorbed by the wires. Finally in this experiment, we record the photons in the area 1' (or 2') and consequently we have at the same time the path information. Importantly, following Afshar *et al*. we here only consider one image spot 1' or 2' (since each photon impinges one only one of these two regions) and we deduce therefore $K = 1$. Together with the interference visibility $V = 1$ this implies

$$K^2 + V^2 = 2, \tag{16}$$

in complete contradiction with the bound given by equation (6).

In the previous analysis we only considered the infinitely thin wires to simplify the discussion. Actually, this is however the only experimental configuration in which the Afshar experiment is easily analyzable since it is only in such case that the duality relation can be defined. Indeed, scattering by the wire always results into complicated diffraction pattern in the image plane and the simple mathematical derivation [25–28] leading to equations 2, 5, and 6 is not possible. We will then continue to consider the idealized case of the infinitely thin wires in the rest of the paper since it is this ideal limit that the authors of [1]wanted obviously to reach.

3 The rebuttal: Inference and Complementarity

3.1 Duality again

There are several reasons why the analysis by Afshar *et al.* actually fails. First, from a mathematical point of view it is not consistent to write $K^2 + V^2 = 2$. Indeed, in all the experiments previously discussed (excluding the Afshar experiments) it was necessary to consider statistics on all the recorded photons in order to observe either the interference or the path information (in the case were entanglement was involved only the photons tagged by $|\lambda\rangle$ have to be considered). Same here, if one consider all the detected photons one will deduce $K = 0$ and equation (6) will be respected. Actually, this results directly from the experimental method considered by the authors of [1]. Indeed, if somebody is accepting the existence of an interference pattern he or she needs to know the complete distribution 1' *and* 2' recorded in the image plane. This is necessary in order to deduce that the wire grid didn't caused any disturbances on the propagation. Indeed, the disturbance could have no consequence in 1' but yet have some effects in 2'. Consequently, ignoring 2' does not allow us to deduce that the experiment with the grid is interaction-free. For this reason, it is unjustified to write $K = 1$, that is to consider only one half of the detected photon population, while we actually need both pinhole images to deduce the value of V (this is also in agreement with the obvious fact that an interference pattern requires the two apertures 1 and 2 opened for its existence).

There is an other equivalent way to see why the choice $K = 0$ is the only one possible. Indeed, having measured in the image plane the two distributions 1' and 2' with intensity $|C_1|^2$ and $|C_2|^2$ we can, by applying the laws of optics, propagate backward in time the two converging beams until the interference plane (this was done by Afshar *et al.*). In this plane equation (4) and (5), which are a direct consequence of these above mentioned optical laws, are of course valid. Since we have $|C_1|^2 = |C_2|^2$, we deduce (from equations (2) and (5)) that $K = 0$ and $V = 1$ in full agreement with the duality relation (6). It is important to remark that since the phase of C_1 and C_2 are not know from the destructive measurements in the image plane, we cannot extrapolate the value of $\chi = \arg(C_1) - \arg(C_2)$. However, the presence of the grid give us access to this missing information since it provides the points where $I(k) = 0$ (for example if $I(\pi/d) = 0$ then $\chi = 2\pi \cdot N$ with N =0, 1, ...). We can thus define completely the variable V and χ without recording any photon in the Fourier plane. It is clear, that this would be impossible if the duality condition $K^2 + V^2 = 1$ was not true since this relation is actually a direct consequence of the law of optics used in our derivations as well as in the one by Afshar *et al.*.

To summarize the present discussion, we showed that Afshar *et al.* reasoning is obscured by a misleading interpretation of the duality relation given by equation (6). We however think that this problem is not so fundamental for the discussion of the experiment. Actually, we can restate the complete reasoning without making any reference to this illusory violation of equation (6). After doing this we think that the error in the deductions by Afshar *et al.* should become very clear. Let then restate the story:

A) First, we record individuals photons in the regions 1' and 2'. We can then keep a track or a list of each detection event, so that, for each photon, we can define its 'path' information. However, this individual property of each photon is not entering in conflict with the statistical behavior, which in the limit of large number, give us the two narrow distribution in 1' and 2'. That is, the value $K = 0$ is not in conflict with the existence of a which-path information associated with each photon. This situation differs strongly from usual which-path experiments in which the path detection, or tagging, is done *before* the interference plane. As we explained before in these experiments the value $K = 1$ was a necessary consequence of the preselection procedure done on the photon population. This point also means that we have to be very prudent when we use the duality relation in experimental situations different from the ones for which a consensus has already been obtained.

B) Second, we apply the laws of optics backward in time to deduce the value of the visibility V. Inferring the validity of such optical laws we can even reconstruct completely the interference profile thanks to the presence of wire grid.

C) Finally, we can check that indeed $K^2+V^2 = 1$ in agreement with the duality relation.

Having elucidated the role of the duality relation, the question that we have still to answer is what are the implications of this experiment for complementarity. What has indeed been shown by Afshar *et al.* is that each photon detected in the image plane is associated with a wave behavior since none of them crossed the wires. Using the laws of optics backward in time allow us to deduce the precise shape of intensity profile in the back focal plane but this is a theoretical inference and actually not a measurement. We will now show that this is the weak point.

3.2 Classical versus quantum inferences

In classical physics, such an inference (i.e., concerning interference) is of no consequence since we can always, at least in principle, imagine a test particle or detector to check the validity of our assumptions concerning the system. However, in quantum mechanics we are dealing with highly sensitive systems and this modify the rules of the game.

In quantum mechanics it is common to say that the wave function represents the catalog of all the potentiality accessible to the system. Due to the very nature of this theory there are however some (complementary) pages which can not be read at the same time without contradictions. In the Afshar experiment, we do not have indeed the slightest experimental proof that the observed photons did participate to the "cos"

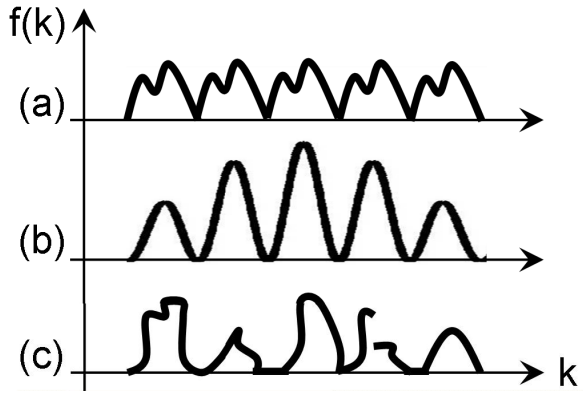


Fig. 2: Different possible intensity profiles in the Fourier plane. Each profile $f(k)$ obeys to the condition $f(k) = 0$ on the wires. (a) A continuous periodic function. (b) The diffractive interference profile predicted by quantum mechanics. (c) A discontinuous profile intensity. Each profile is 'apriori' equiprobable for an observer which has no knowledge in optics and quantum mechanics.

interference pattern given by equations (3) and (4). Furthermore, by detecting the photons in the image plane, we only know from the experiment that the photons never crossed the wires but this is not sufficient to rebuild objectively the complete interference pattern.

We can go further in this direction by using information theory. Indeed, from the point of view of the information theory of Gibbs [33], Shannon [34], and Jaynes [35], every interference patterns, such that $I(k) = 0$ on the wires, are equiprobable (see Fig. 2). However, there are an infinity of such profiles, so that our information is rather poor. More precisely, let write $\rho[f(k)]$ the functional giving the density of probability associated with the apriori likelihood of having the interference profile $f(k)$ located in an infinitely small (functional) volume $\mathcal{D}[f(k)]$. We write $\Sigma[f(x)]$ the space of all this interference profiles obeying to the condition $f(k) = 0$ on the wires. We have thus $\rho[f(k)] = 1/\Sigma$ (equiprobability) for the function f contained in Σ, and $\rho[f(k)] = 0$ for the function outside Σ (that are functions which do not satisfy the requirements $f(k) = 0$ on the wires). The Shannon entropy [33–35] $S[f(x)]$ associated with this distribution is given by

$$S[f(x)] = -\int_{(\Sigma)} \mathcal{D}[f(k)]\rho[f(k)] \ln(\rho[f(k)]) = \ln(\Sigma[f(k)]) \to +\infty, \qquad (17)$$

which expresses our absence of objective knowledge concerning $f(k)$. In this reasoning, we used the concept of probability taken in the Bayesian sense, that is in the sense of decision-maker theory used for example by poker players. For an observer which do not have any idea concerning quantum mechanics and the laws of optics, this equiprobability is the most reasonable guess if he wants only to consider the photons he actually detected. Of course, by considering a different experiment, in which the photons are recorded in the Fourier plane, the observer might realize what is actually the interference pattern. However (and this is essential for understanding the apparent paradox discussed in reference 1) it will be only possible by considering different recorded photons in full agreement with the principle of complementarity.

Let now summarize a bit our analysis. We deduced that in the experiment discussed in [1] the photons used to *measure* objectively the interference pattern i.e. to calculate the visibility $V = 1$ are not the same than those used to *measure* the distribution in the image plane and calculate the distinguishability $K = 0$. This is strictly the same situation than in the original two-holes experiment already mentioned. It is in that sense that the relationship (6) represents indeed a particular formulation of complementarity [25–28]. Actually (as we already commented before) the value $V = 1$ obtained in [1] does not result from a measurement but from an extrapolation. Indeed, from their negative measurement Afshar et al. recorded objectively $I_{min} = 0$. If we suppose that there is a hidden sinusoidal interference pattern in the plane of the wires we can indeed write

$$V = (I_{max} - I_{min}) / (I_{max} + I_{min}) = I_{max}/I_{max} = 1. \qquad (18)$$

However to prove experimentally that such sinusoidal interference pattern actually exists we must definitively record photons in the rest of the wires plane. This is why the experiment described in [1] does not constitutes a violation of complementarity.

It is finally interesting to remark that similar analysis could be easily done already in the Young two-holes experiment. Indeed, suppose that we record the photon interference fringes after the holes. We can thus measure $V = 1$. However, if we suppose that the sinusoidal oscillation of the intensity results from the linear superposition of waves coming from holes 1 and 2 then from equation 5 we deduce $|C_1|^2 + |C_2|^2 - 2|C_1||C_2| = 0$ i. e., $|C_1| = |C_2|$. From equation 2 this implies $K = 0$. Reasoning like Afshar *et al.* we could be tempted to see once again a violation of complementarity since we deduced the distinguishability without disturbing the fringes! However, we think that our previous analysis sufficiently clarified the problem so that paradoxes of that kind are now naturally solved without supplementary comments.

3.3 The objectivity of trajectory in quantum mechanics

At the end of section 2.1 we shortly pointed that the concept of trajectory is a key issue in the analysis of the experiment reported in reference 1. This was also at the core of most commentaries (e.g. references [6–14]) concerning the work by Afshar *et al.*. As a corollary to the previous analysis we will now make a brief comment concerning the concept of path and trajectory in quantum mechanics since we think that a lot of confusion surrounds this problem. This is also important because Afshar *et al.* claimed not only that they can

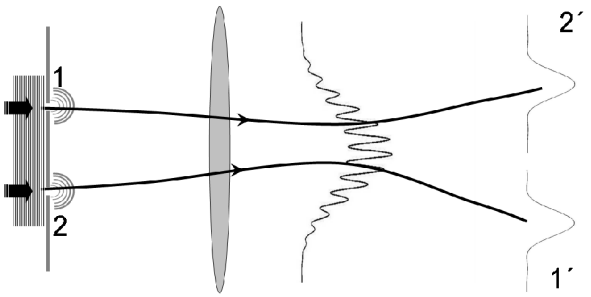


Fig. 3: Illustration of the counterintuitive paths followed by photons if we accept the ontological interpretation given by de Broglie and Bohm. The photons coming from aperture 1 or 2 reach the 'wrong' detector 2' or 1'.

circumvent complementarity but that additionally they determine the *path* chosen by the particle. Following here an intuitive assumption they accepted that with the two pinholes open a photon trajectory (if trajectory there is) connects necessarily a pinhole to its optical image like it is in geometrical optics. They called that intuition (probably in analogy with what occurs in classical physics) a 'consequence of momentum conservation'. However, the meaning of momentum and trajectory is not the same in quantum and classical mechanics. Actually, as it was realized by several physicists the connection 1 to 1' and 2 to 2' is a strong hypothesis which depends of our model of (hidden) reality and which can not in general be experimentally tested (read for example [29, 36]).

Actually nothing in this experiment with two holes forbids a photon coming from one pinhole to go in the *wrong* detector associated with the second pinhole. This is the case for example in the *hidden variable* theory of de Broglie-Bohm in which every photons coming from the aperture 1 (respectively 2) is reaching the wrong image spot 2'(respectively 1') [29, 36] as shown in figure 3. This is counter intuitive but not in contradiction with experiments since we can not objectively test such hidden variable model [36]. In particular closing one pinhole will define unambiguously the path followed by the particle. However this is a different experiment and the model shows that the trajectories are modified (in general non locally) by the experimental context. The very existence of a model like the one of de Broglie and Bohm demonstrates clearly that in the (hidden) quantum reality a trajectory could depend of the complete context of the experiment. For this reason we must be very prudent and conservative when we interpret an experiment: Looking the image of a pinhole recorded in a statistical way by a cascade of photon will not tell us from which pinhole an individual photon come from but only how many photons crossed this pinhole. In counterpart of course we can not see the fringes and the complementarity principle of Bohr will be, as in every quantum experiment, naturally respected. It is thus in general dangerous to speak unambiguously of a which-path experiment and this should preferably be avoided from every discussions limited to empirical facts. As claimed by Bohr the best empirical choice is in such conditions to accept that *it is wrong to think that the task of physics is to find out how Nature is. Physics concerns what we can say about Nature* [4].

4 Conclusion

To conclude, in spite of some claims we still need at least two complementary experiments in order to exploit the totality of the phenomenon in Young-like interferometers. Actually, as pointed out originally by Bohr, we can not use information associated with a same photon event to reconstruct in a statistical way (i.e. by a accumulation of such events) the two complementary distributions of photons in the image plane of the lens and in the interference plane. The presence of the wires inserted in reference 1 does not change anything to this fact since the information obtained by adding the wires is too weak and not sufficient to rebuild objectively (i. e. , unambiguously from experimental data) the whole interference pattern. The reasoning of Afshar *et al.* is therefore circular and the experiment is finally in complete agreement with the principle of complementarity.

Submitted on September 12, 2010 / Accepted on September 15, 2010

References

1. Afshar S., Flores E., McDonalds K. F., Knoesel E. Paradox in wave-particle duality. *Foundations of Physics*, 2007, v. 37, 295-–305.
2. Afshar S. Violation of Bohr's complementarity: one slit or both? *AIP Conference Proceedings*, 2006 v. 810, 294-–299.
3. Bohr N. Discussions with Einstein on epistemological problems in atomic physics. Albert Einstein philosopher-scientist, edited by P. A. Schilpp, The library of living philosophers, Evanston, 1949, pp. 200–241.
4. Bohr N. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 1935, v. 48, 696–702.
5. Drezet A. Complementarity and Afshar's experiment. arXiv:quant-ph/0508091v3.
6. Kastner R. Why the Afshar experiment does not refute complementarity? *Studies in history and philosophy of modern physics*, 2005 v. 36, 649–658.
7. Kastner R. On Visibility in the Afshar Two-Slit Experiment. *Foundations of Physics*, 2009, v. 39, 1139–1144.
8. Steuernagel O. Afshar's Experiment Does Not Show a Violation of Complementarity. *Foundations of Physics*, 2007, v. 37, 1370–1385.
9. Qureshi T. Complementarity and the Afshar Experiment. arXiv:quant-ph/0701109v2.
10. Flores E. V. Reply to Comments of Steuernagel on the Afshar's Experiment. *Foundations of Physics*, 2008, v. 38, 778–781.
11. Flores E. V., Knoesel E. Why Kastner analysis does not apply to a modified Afshar experiment. arXiv:quant-ph/0702210.
12. Jacques V. et al. Illustration of quantum complementarity using single photons interfering on a grating. *New Journal of Physics*, 2008, v. 10, 123009.
13. Georgiev D. D. Single photon experiments and quantum complementarity. *Progress in Physics*, 2007, v. 2, 97–103; Ibid 2007, v. 3, 28.

14. Unruh W. G. Comment on "single photon experiments and quantum complementarity" by D. Georgiev. *Progress in Physics*, 2007, v. 3, 27–27. See also: http://axion.physics.ubc.ca/rebel.html
15. Feynman R. P., Leighton R., Sand M. The Feynman Lectures on Physics. Vol. 3, Adisson Wesley, Reading, 1965.
16. Zeilinger A. Experiment and the foundations of quantum physics. *Review of Modern Physics*, 1999, v. 71, S288–S297.
17. Scully M. O., Englert B. G., Walther H. Quantum optical tests of complementarity. *Nature*, 1991, v. 351, 111–116.
18. Drezet A., Hohenau H., Krenn J. R. Heisenberg optical near-field microscope. *Physical Review A*, 2006, v. 73, 013402.
19. Drezet A., Hohenau H., Krenn J. R. Momentum transfer for momentum transfer-free which-path experiments. *Physical Review A*, 2006, v. 73, 062112.
20. de Broglie L. Ondes et Mouvements. Gauthier-Villars, Paris, 1926.
21. Bohm D. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. Part 1 and 2. *Physical Review* 1952, v. 85, 166–179 and 180–193.
22. Greene B. The fabric of the cosmos. Alfred A. Knopf, New York 2004.
23. Wheeler J. A. Mathematical Foundations of Quantum Physics. A. R. Marlow (Editor), Academic, NewYork, 1978.
24. In the actual experimental setup considered in [1] the lens is located far away from the two pinholes so that it is in practice equivalent to observe the fringes in front of the lens or in its back focal plane.
25. Englert B. G. Fringe Visibility and Which-Way Information: An Inequality. *Physical Review Letters*, 1996, v. 77, 2154–2157.
26. Greenberger D. M., Yasin A. Simultaneous wave and particle knowledge in a neutron interferometer. *Physics Letters A*, 1988, v. 128, 391–394.
27. Wooters W. K., Zurek W. H. Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle. *Physical Review D*, 1979, v. 19, 473–484.
28. Jaeger G., Shimony A., Vaidman L. Two interferometric complementarities. *Physical Review A*, 1995, v. 51, 54–67.
29. Englert B.-G., Scully M. O., Süssmann G., Walther H. Surrealistic Bohm trajectories. *Zeitschrift für Naturforschung A*, 1992, v. 47, 1175–1186.
30. Vaidman L. The reality in bohmian quantum mechanics or can you kill with an empty wave bullet? *Foundations of Physics*, 2005, v. 35, 299–312.
31. Bohm D. J., Dewdney C., Hiley B. H. A quantum potential approach to the Wheeler delayed-choice experiment. *Nature*, 1985, v. 315, 294–297.
32. Bartell L. S. Complementarity in the double-slit experiment: On simple realizable systems for observing intermediate particle-wave behavior. *Physical Review D*, 1980, v. 21, 1698–1699.
33. Gibbs J. W. Elementary Principles in Statiscal Mechanics. Longmans Green and Company, NewYork, 1928.
34. Shannon C. E., Weaver W. The mathematical theory of communications. University of Illinois Press, Urbana,1949.
35. Jaynes E. T. Information theory and statistical mechanics. *Physical Review*, 1957 v. 106, 620–630.
36. Hiley B. J. and Callaghan R. E. What is Erased in the Quantum Erasure? *Foundations of Physics*, 2007, v.36, 1869-1883.

Noise and Fano-factor Control in AC-Driven Aharonov-Casher Ring

Walid A. Zein*, Nabil A. Ibrahim†, and Adel H. Phillips*

*Faculty of Engineering, Ain-Shams University, Cairo, Egypt
†Higher Technological Institute, Ramadan Tenth City, Egypt
E-mail: adel_phillips@yahoo.com

The spin dependent current and Fano factor of Aharonov-Casher semiconducting ring is investigated under the effect of microwave, infrared, ultraviolet radiation and magnetic field. Both the average current and the transport noise (Fano factor) characteristics are expressed in terms of the tunneling probability for the respective scattering channels. For spin transport induced by microwave and infrared radiation, a random oscillatory behavior of the Fano factor is observed. These oscillations are due to constructive and destructive spin interference effects. While for the case of ultraviolet radiation, the Fano factor becomes constant. This is due to that the oscillations has been washed out by phase averaging (i.e. ensemble dephasing) over the spin transport channels. The present investigation is very important for quantum computing and information processing.

1 Introduction

The field of spintronics is devoted to create, store, manipulate at a given location, and transport coherent electron spin states through dilute magnetic semiconductors and conventional semiconductor heterostructure [1]. The two principle challenges for new generation of spintronics devices are efficient injection of spins into various semiconductor nanostructures and coherent control of spin. In particular, preserving spin coherence, which enables coherent superpositions of states $a|\uparrow\rangle + b|\downarrow\rangle$ and corresponding quantum-interference effects, is essential for both quantum computing with spin-based qubits [2, 3]. The electrical control of spin via Rashba spin-orbit coupling [4], which arises due to inversion asymmetry of the confining electric potential for tow-dimensional electron gas (2DEG), is very important physical parameter when dealing with semiconductor spintronics. The pursuit of fundamental spin interference effects, as well as spin transistors with unpolarized charge currents [3, 5–10] has generated considerable interest to demonstrate the Aharonov-Casher effect via transport experiments in spin-orbit coupled semiconductor nanostructures [7, 11].

The ballistic spin-resolved shot noise and consequently Fano factor in Aharonov-Casher semiconducting ring is investigated in the present paper. The effects of both electromagnetic field of wide range of frequencies and magnetic field are taken into consideration.

2 Theoretical Formulation

It is well known that shot noise and consequently Fano factor is a powerful quantity to give information about controlling decoherence of spin dependent phenomena [12, 13]. So we shall deduce an expression for both shot noise and Fano factor for spintronic device considered in the paper [10]. This device is modeled as follows: Aharonov-Casher interferometer ring in which a semiconductor quantum dot is embedded

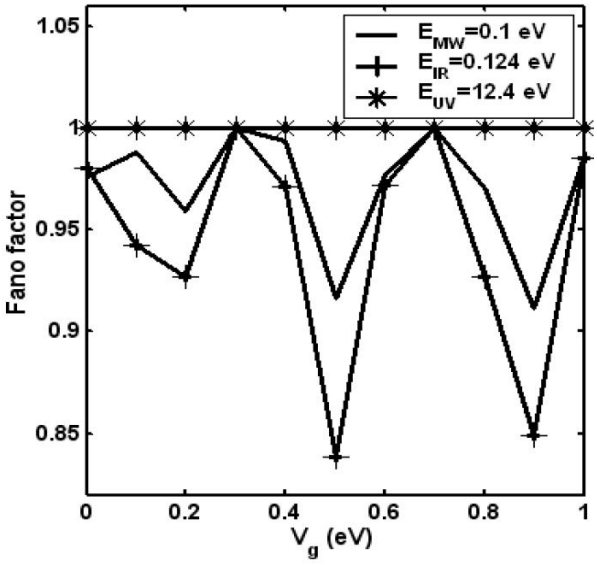


Fig. 1: The variation of Fano factor with gate voltage at different photon energies.

in one arm of the ring. The form of the confining potential is modulated by an external gate electrode allowing for direct control of the electron spin-orbit coupling. The effect of electromagnetic field of wide range of frequencies (microwave, infrared, ultraviolet) is taken into consideration.

The spin dependent shot noise $S_{\alpha\beta}^{\sigma\sigma'}(t-t')$ is expressed in terms of the spin resolved currents $I(\uparrow)$, and $I(\downarrow)$ due to the flow of spin-up $\uparrow$ and spin-down $\downarrow$ electrons through the terminals of the present device [14] as

$$S_{\alpha\beta}^{\sigma\sigma'}(t-t') = \frac{1}{2}\left\langle \delta\hat{I}_{\alpha}^{\sigma}(t)\,\delta\hat{I}_{\beta}^{\sigma'}(t') + \delta\hat{I}_{\beta}^{\sigma'}(t')\,\delta\hat{I}_{\alpha}^{\sigma}(t)\right\rangle \quad (1)$$

where $\hat{I}_{\alpha}^{\sigma}(t)$ is the quantum mechanical operator of the spin resolved ($\sigma \Rightarrow \uparrow, \downarrow$) current in left lead α, $\hat{I}_{\beta}^{\sigma'}(t')$ is the same

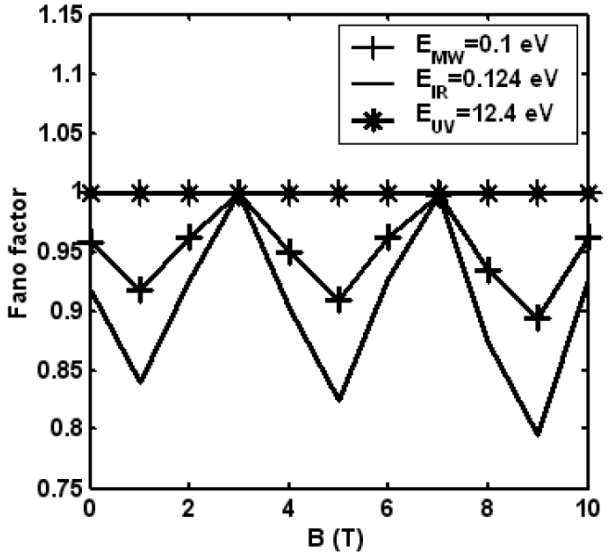


Fig. 2: The variation of Fano factor with magnetic field at different photon energies.

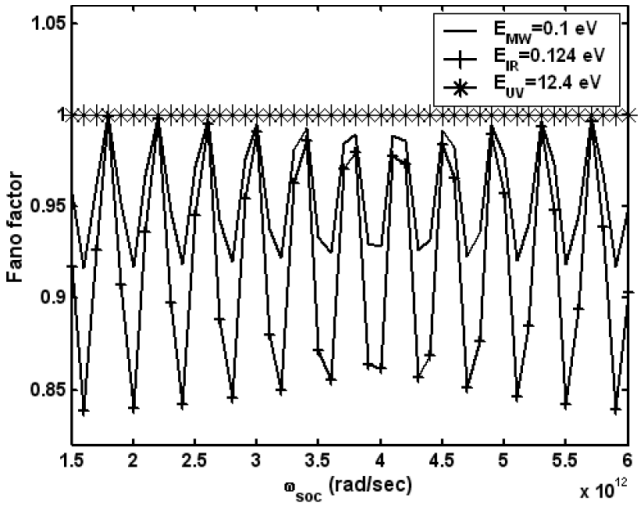


Fig. 3: The variation of Fano factor with frequency ω_{Soc} at different photon energies.

definition of $\hat{I}^{\sigma}_{\alpha}(t)$, but for the right lead β. In Eq. (1), the parameter $\delta\hat{I}^{\sigma}_{\alpha}(t)$ represents the current fluctuation operator at time t in the left lead α with spin state σ (up or down) and is given by

$$\delta\hat{I}^{\sigma}_{\alpha}(t) = \hat{I}^{\sigma}_{\alpha}(t) - \left\langle \hat{I}^{\sigma}_{\alpha}(t) \right\rangle \tag{2}$$

where $\langle --\rangle$ denotes an ensemble average. The Fourier transform of Eq.(1), which represents the spin resolved noise power between the left and right terminals of the device, is given by

$$S^{\sigma\sigma'}_{\alpha\beta}(\omega) = 2\int d(t-t')\, e^{-i\omega(t-t')} S^{\sigma\sigma'}_{\alpha\beta}(t-t'). \tag{3}$$

Since the total spin dependent current is given by

$$I_{\alpha} = I^{\uparrow}_{\alpha} + I^{\downarrow}_{\alpha}, \tag{4}$$

the corresponding noise power is expressed as

$$S_{\alpha\beta}(\omega) = S^{\uparrow\uparrow}_{\alpha\beta}(\omega) + S^{\downarrow\downarrow}_{\alpha\beta}(\omega) + S^{\uparrow\downarrow}_{\alpha\beta}(\omega) + S^{\downarrow\uparrow}_{\alpha\beta}(\omega). \tag{5}$$

Now, expressing the spin-resolved current $\hat{I}^{\sigma}_{\alpha}(t)$ in terms of the creation and annihilation operators of the incoming electrons $\hat{a}^{\sigma+}_{\alpha}(E)$, $\hat{a}^{\sigma}_{\alpha}(E')$ and for the outgoing electrons $\hat{b}^{\sigma+}_{\alpha}(E+n\hbar\omega)$, $\hat{b}^{\sigma}_{\alpha}(E'+n\hbar\omega)$ [15], as follows:

$$\hat{I}^{\sigma}_{\alpha}(t) = \frac{e}{h}\sum_{n}\int\int dE dE' e^{i(E-E')t/\hbar} \times \left[\hat{a}^{\sigma+}_{\alpha}(E)\,\hat{a}^{\sigma}_{\alpha}(E') - \hat{b}^{\sigma+}_{\alpha}(E+n\hbar\omega)\,\hat{b}^{\sigma}_{\alpha}(E'+n\hbar\omega)\right]. \tag{6}$$

Now, in order to evaluate the shot noise spectrum $S_{\alpha\beta}(\omega)$ this can be achieved by substituting Eq.(6) into Eq.(1), and using the transmission eigenfunctions [10] through the present spintronic device, we can determine the expectation value [15, 16]. We get an expression for the shot noise spectrum $S_{\alpha\beta}(\omega)$ as follows:

$$S_{\alpha\beta}(\omega) = \frac{2eP_0}{h}\sum_{\sigma}\int_{0}^{\infty} dE \left|\Gamma_{\mu\ with\ photon}(E)\right|^2 \times f_{\alpha FD}(E) \times \left[1 - f_{\beta FD}(E+n\hbar\omega)\right] \tag{7}$$

where $\left|\Gamma_{\mu\ with\ photon}(E)\right|^2$ is the tunneling probability induced by the external photons, and $f_{\alpha FD}(E)$, $f_{\beta FD}(E+n\hbar\omega)$ are the Fermi distribution functions, and P_0 is the Poissonian shot noise spectrum [15].

The tunneling probability $\left|\Gamma_{\mu\ with\ photon}(E)\right|^2$ has been determined previously by the authors [10]

The Fano factor, F, of such mesoscopic device is given by [17]:

$$F = \frac{S_{\alpha\beta}(\omega)}{2eI}. \tag{8}$$

The explicit expression for the Fano factor, F, can be written as, after some algebraic computation of Eqs.(7, 8), [18, 19]:

$$F = \frac{\left[\sum_{n}\sum_{\mu}\left|\Gamma_{\mu\ with\ photon}(E)\right|^2\left(1 - \left|\Gamma_{\mu\ with\ photon}(E)\right|^2\right)\right]}{\sum_{n}\sum_{\mu}\left|\Gamma_{\mu\ with\ photon}(E)\right|^2}. \tag{9}$$

3 Results and Discussion

The Fano factor F Eq.(9) has been computed numerically as a function of the gate voltage V_g magnetic field B and function of the frequency ω_{Soc} due to spin-orbit coupling. These calculations are performed over a wide range of frequencies

of the induced electromagnetic field (microwave, MW, infrared, IR, and ultraviolet, UV). We use the semiconductor heterostructures as *InGaAs/InAlAs* as in the paper [10]. The main features of the present obtained results are:

(1) Fig.1, shows the dependence of Fano factor on the gate voltage V_g at photon energies for microwave, infrared, and ultraviolet. As shown from the figure that, the Fano factor fluctuates between maximum and minimum values for the two cases microwave and infrared irradiation. While for the case of ultraviolet irradiation, the Fano factor is constant and approximately equals $\sim$1.

(2) Fig.2, shows the dependence of Fano factor on the magnetic field B at photon energies for microwave, infrared, and ultraviolet. The trend of this dependence is similar in a quite fair to the trend and behavior of Fig.1.

(3) Fig.3, shows the dependence of Fano factor on the frequency ω_{Soc} associated with the spin-orbit coupling at photon energies for microwave, infrared, and ultraviolet. An oscillatory behavior for this dependence for the two cases microwave and infrared are shown. While for the case of ultraviolet, the Fano factor is constant and approximately equals $\sim$1 as in Figs. 1, 2.

These results might be explained as follows: Computations show that the average current suppression is accompanied by a noise maxima and remarkably low minima (Fano factor). These cases are achieved when the electron spin transport is influenced by both microwave and infrared photons. Such results have been observed previously by the authors [20–22]. The random oscillatory behavior of the Fano factor can be understood as the strength of the spin-orbit coupling is modified by the gate electrode covering the Aharonov-Casher ring to tune constructive and destructive spin interference effect [10]. For the case of the induced ultraviolet radiation, the results show that the Fano factor becomes approximately constant. These results have been observed previously by the authors [23,24]. The constancy of Fano factor might be due to washing out of the oscillations by phase averaging (i.e. ensemble dephasing) over the spin transport channels [23, 24].

We conclude that these phenomena can be used to devise novel spintronic devices with a priori controllable noise levels. The present investigation is very important for quantum computing and quantum information processing.

Submitted on September 9, 2010 / Accepted on September 14, 2010

References

1. Zutic I., Fabian J., Das Sarma S. Spintronics: Fundamentals and Applications. *Review of Modern Physics*, 2004, v. 76, 323–410.
2. Nikolic B. K., Zarbo L. P., Souma S. Imaging Mesoscopic Spin Hall Flow: Spatial Distribution of Local Spin Currents and Spin Densities in and out of Multiterminal Spin-Orbit Coupled Semiconductor Nanostructures. *Physical Review B*, 2006, v. 73, 075303.
3. Fabian J., Matos-Abiaguea A., Ertlera C., Stano P., Zutic I. Semiconductor Spintronics. *Acta Physica Slovaca*, 2007, v. 57, 565–907.
4. Rashba E. I. Electron Spin Operation by Electric Fields: Spin Dynamics and Spin Injection. *Physica E*, 2004, v. 20, 189–195.
5. Nitta J., Meijer F. E., Takayanagi H. Spin Interference Device. *Applied Physics Letters*, 1999, v. 75, 695–697.
6. Nitta J., Bergsten T. Time Reversal Aharonov-Casher Effect Using Rashba Spin-Orbit Interaction. *New Journal of Physics*, 2007, v.9, 341–352.
7. Frustaglia D., Richter K. Spin Interference Effects in Ring Conductors Subject to Rashba Coupling. *Physical Review B*, 2004, v.69, 235310.
8. Zein W. A., Phillips A. H., Omar O. A. Quantum Spin Transport in Mesoscopic Interferometer. *Progress in Physics*, 2007, v. 4, 18–21.
9. Zein W. A., Phillips A. H., Omar O. A. Spin Coherent Transport in Mesoscopic Interference Device. *NANO*, 2007, v. 2, no. 6, 389–392.
10. Zein W. A., Ibrahim N. A., Phillips A. H. Spin Dependent Transport through Aharonov-Casher Ring Irradiated by an Electromagnetic Field. *Progress in Physics*, 2010, v. 4, 78–81.
11. Konig M., Tschetschetkin A., Hankiewicz E. M., Sinova J., Hock V., Daumer V., Schafer M., Beacker C. R., Buhmann H., Molenkamp L. W. Direct Observation of the Aharonov-Casher Phase. *Physical Review Letters*, 2006, v. 96, 076804.
12. Awschalom D. D., Flatte M. E. Challenges for Semiconductor Spintronics. *Nature Physics*, 2007, v. 3, 153–159.
13. Sukhorukov E. V., Burkard G., Loss D. Noise of a Quantum dot System in the Cotunneling Regime. *Physical Review B*, 2001, v. 63, 125315.
14. Sauret O., Feinberg D. Spin-Current Shot Noise as a probe of Interactions in Mesoscopic Systems. *Physical Review Letters*, 2004, v. 92, 106601.
15. Mina A. N., Phillips A. H. Frequency Resolved Detection over a Large Frequency Range of the Fluctuations in an Array of Quantum Dots. *Progress in Physics*, 2006, v. 4, 11–17.
16. Beenakker C. W. J., Buttiker M. Suppression of Shot Noise in Metallic Diffusive Conductors. *Physical Review B*, 1992, v. 46, R1889.
17. Blanter Ya. M., Buttiker M., Shot Noise in Mesoscopic Conductors. *Physics Reports*, 2000, v. 336, 1–166.
18. Dragomirova R. L., Nikolic B. K. Shot Noise of Spin Polarized Charge Currents as a Probe of Spin Coherence in Spin-Orbit Coupled Nanostructures. *Physical Review B*, 2007, v. 75, 085328.
19. Liang-Zhong L., Rui Z., Wen-Ji D. Shot Noise in Aharonov-Casher Rings. *Chinese Physics Letters*, 2010, v. 27, no. 6, 067306.
20. Camalet S., Lehmann J., Kohler S., Hanggi P. Current Noise in ac-driven Nanoscale Conductors. *Physical Review Letters*, 2003, v. 90, 210602.
21. Camalet S., Kohler S., Hanggi P. Shot Noise Control in AC-driven Nanoscale Conductors. *Physical Review B*, 2004, v. 70, 155326.
22. Sanchez R., Kohler S., Platero G. Spin Correlation in Spin Blockade. *New Journal of Physics*, 2008, v. 10, 115013.
23. Souma S., Nikolic B. K. Modulating unpolarized Current in Quantum Spintronics: Visibility of Spin Interfering Effects in Multichannel Aharonov-Casher Mesoscopic Rings. *Physical Review B*, 2004, v. 70, 195346.
24. Padurariu C., Amin A. F., Kleinekathofer U. Laser-Assisted Electron Transport in Nanoscale Devices, in: Radons G., Rumpf B., Schuster H. G (Editors), Nonlinear Dynamics of Nanosystems, Wiley-VCH, 2009.

Smarandache's Minimum Theorem in the Einstein Relativistic Velocity Model of Hyperbolic Geometry

Cătălin Barbu
"Vasile Alecsandri" College — Bacău, str. Vasile Alecsandri, nr.37, 600011, Bacău, Romania. E-mail: kafka_mate@yahoo.com.

In this note, we present a proof to the Smarandache's Minimum Theorem in the Einstein Relativistic Velocity Model of Hyperbolic Geometry.

1 Introduction

Hyperbolic Geometry appeared in the first half of the 19^{th} century as an attempt to understand Euclid's axiomatic basis of Geometry. It is also known as a type of non-Euclidean Geometry, being in many respects similar to Euclidean Geometry. Hyperbolic Geometry includes similar concepts as distance and angle. Both these geometries have many results in common but many are different. There are known many models for Hyperbolic Geometry, such as: Poincaré disc model, Poincaré half-plane, Klein model, Einstein relativistic velocity model, etc. Here, in this study, we give hyperbolic version of Smarandache minimum theorem in the Einstein relativistic velocity model of hyperbolic geometry. The well-known Smarandache minimum theorem states that if ABC is a triangle and AA', BB', CC' are concurrent cevians at P, then

$$\frac{PA}{PA'} \cdot \frac{PB}{PB'} \cdot \frac{PC}{PC'} \geq 8$$

and

$$\frac{PA}{PA'} + \frac{PB}{PB'} + \frac{PC}{PC'} \geq 6$$

(see [1]).

Let D denote the complex unit disc in complex z-plane, i.e.

$$D = \{z \in \mathbb{C} : |z| < 1\}.$$

The most general Möbius transformation of D is

$$z \to e^{i\theta} \frac{z_0 + z}{1 + \overline{z_0} z} = e^{i\theta}(z_0 \oplus z),$$

which induces the Möbius addition $\oplus$ in D, allowing the Möbius transformation of the disc to be viewed as a Möbius left gyrotranslation

$$z \to z_0 \oplus z = \frac{z_0 + z}{1 + \overline{z_0} z}$$

followed by a rotation. Here $\theta \in \mathbb{R}$ is a real number, $z, z_0 \in D$, and $\overline{z_0}$ is the complex conjugate of z_0. Let $Aut(D, \oplus)$ be the automorphism group of the grupoid $(D, \oplus)$. If we define

$$gyr : D \times D \to Aut(D, \oplus), gyr[a, b] = \frac{a \oplus b}{b \oplus a} = \frac{1 + a\overline{b}}{1 + \overline{a}b},$$

then is true gyrocommutative law

$$a \oplus b = gyr[a, b](b \oplus a).$$

A gyrovector space $(G, \oplus, \otimes)$ is a gyrocommutative gyrogroup $(G, \oplus)$ that obeys the following axioms:
(1) $gyr[\mathbf{u}, \mathbf{v}]\mathbf{a} \cdot gyr[\mathbf{u}, \mathbf{v}]\mathbf{b} = \mathbf{a} \cdot \mathbf{b}$ for all points $\mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{v} \in G$.
(2) G admits a scalar multiplication, $\otimes$, possessing the following properties. For all real numbers $r, r_1, r_2 \in \mathbb{R}$ and all points $\mathbf{a} \in G$:
(G1) $1 \otimes \mathbf{a} = \mathbf{a}$
(G2) $(r_1 + r_2) \otimes \mathbf{a} = r_1 \otimes \mathbf{a} \oplus r_2 \otimes \mathbf{a}$
(G3) $(r_1 r_2) \otimes \mathbf{a} = r_1 \otimes (r_2 \otimes \mathbf{a})$
(G4) $\frac{|r| \otimes \mathbf{a}}{\|r \otimes \mathbf{a}\|} = \frac{\mathbf{a}}{\|\mathbf{a}\|}$
(G5) $gyr[\mathbf{u}, \mathbf{v}](r \otimes \mathbf{a}) = r \otimes gyr[\mathbf{u}, \mathbf{v}]\mathbf{a}$
(G6) $gyr[r_1 \otimes \mathbf{v}, r_1 \otimes \mathbf{v}] = 1$
(3) Real vector space structure $(\|G\|, \oplus, \otimes)$ for the set $\|G\|$ of onedimensional "vectors"

$$\|G\| = \{\pm \|\mathbf{a}\| : \mathbf{a} \in G\} \subset \mathbb{R}$$

with vector addition $\oplus$ and scalar multiplication $\otimes$, such that for all $r \in \mathbb{R}$ and $\mathbf{a}, \mathbf{b} \in G$,
(G7) $\|r \otimes \mathbf{a}\| = |r| \otimes \|\mathbf{a}\|$
(G8) $\|\mathbf{a} \oplus \mathbf{b}\| \leqslant \|\mathbf{a}\| \oplus \|\mathbf{b}\|$

Theorem 1. ***(Ceva's theorem for hyperbolic triangles).*** *If M is a point not on any side of an gyrotriangle ABC in a gyrovector space $(V_s, \oplus, \otimes)$, such that AM and BC meet in A', BM and CA meet in B', and CM and AB meet in C', then*

$$\frac{\gamma_{|AC'|}\,|AC'|}{\gamma_{|BC'|}\,|BC'|} \cdot \frac{\gamma_{|BA'|}\,|BA'|}{\gamma_{|CA'|}\,|CA'|} \cdot \frac{\gamma_{|CB'|}\,|CB'|}{\gamma_{|AB'|}\,|AB'|} = 1,$$

where $\gamma_{\mathbf{v}} = \frac{1}{\sqrt{1 - \frac{\|\mathbf{v}\|^2}{s^2}}}$.

(See [2, p. 564].) For further details we refer to the recent book of A.Ungar [3].

Theorem 2. ***(Van Aubel's theorem in hyperbolic geometry).*** *If the point P does lie on any side of the hyperbolic triangle ABC, and BC meets AP in D, CA meets BP in E, and AB meets CP in F, then*

$$\frac{\gamma_{|AP|}\,|AP|}{\gamma_{|PD|}\,|PD|} = \frac{\gamma_{|BC|}\,|BC|}{2}\left(\frac{\gamma_{|AE|}\,|AE|}{\gamma_{|EC|}\,|EC|} \cdot \frac{1}{\gamma_{|BD|}\,|BD|}\right) + \frac{\gamma_{|BC|}\,|BC|}{2}\left(\frac{\gamma_{|FA|}\,|FA|}{\gamma_{|FB|}\,|FB|} \cdot \frac{1}{\gamma_{|CD|}\,|CD|}\right).$$

(See [4].)

2 Main result

In this section, we prove Smarandache's minimum theorem in the Einstein relativistic velocity model of hyperbolic geometry.

Theorem 3. *If ABC is a gyrotriangle and AA', BB', CC' are concurrent cevians at P, then*

$$\frac{\gamma_{|AP|}\,|AP|}{\gamma_{|PA'|}\,|PA'|}\cdot\frac{\gamma_{|BP|}\,|BP|}{\gamma_{|PB'|}\,|PB'|}\cdot\frac{\gamma_{|CP|}\,|CP|}{\gamma_{|PC'|}\,|PC'|}\geqslant 1,$$

and

$$\frac{\gamma_{|AP|}\,|AP|}{\gamma_{|PA'|}\,|PA'|}+\frac{\gamma_{|BP|}\,|BP|}{\gamma_{|PB'|}\,|PB'|}+\frac{\gamma_{|CP|}\,|CP|}{\gamma_{|PC'|}\,|PC'|}\geqslant 3.$$

Proof. We set

$$|A'C|=a_1,|BA'|=a_2,|B'A|=b_1,$$

$$|B'C|=b_2,|C'B|=c_1,|C'A|=c_2,$$

$$\frac{\gamma_{|AP|}\,|AP|}{\gamma_{|PA'|}\,|PA'|}\cdot\frac{\gamma_{|BP|}\,|BP|}{\gamma_{|PB'|}\,|PB'|}\cdot\frac{\gamma_{|CP|}\,|CP|}{\gamma_{|PC'|}\,|PC'|}=P,$$

$$\frac{\gamma_{|AP|}\,|AP|}{\gamma_{|PA'|}\,|PA'|}+\frac{\gamma_{|BP|}\,|BP|}{\gamma_{|PB'|}\,|PB'|}+\frac{\gamma_{|CP|}\,|CP|}{\gamma_{|PC'|}\,|PC'|}=S.$$

If we use the Van Aubel's theorem in the gyrotriangle ABC (See Theorem 2), then

$$\frac{\gamma_{|AP|}|AP|}{\gamma_{|PA'|}|PA'|}=\frac{\gamma_{|BC|}|BC|}{2}\left(\frac{\gamma_{|AB'|}|AB'|}{\gamma_{|CB'|}|CB'|}\cdot\frac{1}{\gamma_{|BA'|}|BA'|}\right)+\frac{\gamma_{|BC|}|BC|}{2}\left(\frac{\gamma_{|AC'|}|AC'|}{\gamma_{|BC'|}|BC'|}\cdot\frac{1}{\gamma_{|CA'|}|CA'|}\right)$$

$$=\frac{\gamma_a a}{2}\left[\frac{\gamma_{b_1}b_1}{\gamma_{b_2}b_2}\cdot\frac{1}{\gamma_{a_2}a_2}+\frac{\gamma_{c_2}c_2}{\gamma_{c_1}c_1}\cdot\frac{1}{\gamma_{a_1}a_1}\right],\qquad(1)$$

and

$$\frac{\gamma_{|BP|}|BP|}{\gamma_{|PB'|}|PB'|}=\frac{\gamma_{|CA|}|CA|}{2}\left(\frac{\gamma_{|BC'|}|BC'|}{\gamma_{|AC'|}|AC'|}\cdot\frac{1}{\gamma_{|CB'|}|CB'|}\right)+\frac{\gamma_{|CA|}|CA|}{2}\left(\frac{\gamma_{|BA'|}|BA'|}{\gamma_{|CA'|}|CA'|}\cdot\frac{1}{\gamma_{|AB'|}|AB'|}\right)$$

$$=\frac{\gamma_b b}{2}\left[\frac{\gamma_{c_1}c_1}{\gamma_{c_2}c_2}\cdot\frac{1}{\gamma_{b_2}b_2}+\frac{\gamma_{a_2}a_2}{\gamma_{a_1}a_1}\cdot\frac{1}{\gamma_{b_1}b_1}\right],\qquad(2)$$

and

$$\frac{\gamma_{|CP|}|CP|}{\gamma_{|PC'|}|PC'|}=\frac{\gamma_{|AB|}|AB|}{2}\left(\frac{\gamma_{|CA'|}|CA'|}{\gamma_{|BA'|}|BA'|}\cdot\frac{1}{\gamma_{|AC'|}|AC'|}\right)+\frac{\gamma_{|AB|}|AB|}{2}\left(\frac{\gamma_{|CB'|}|CB'|}{\gamma_{|AB'|}|AB'|}\cdot\frac{1}{\gamma_{|BC'|}|BC'|}\right)$$

$$=\frac{\gamma_c c}{2}\left(\frac{\gamma_{a_1}a_1}{\gamma_{a_2}a_2}\cdot\frac{1}{\gamma_{c_2}c_2}+\frac{\gamma_{b_2}b_2}{\gamma_{b_1}b_1}\cdot\frac{1}{\gamma_{c_1}c_1}\right).\qquad(3)$$

If we use the Ceva's theorem in the gyrotriangle ABC (See Theorem 1), we have

$$\frac{\gamma_{|CA'|}\,|CA'|}{\gamma_{|BA'|}\,|BA'|}\cdot\frac{\gamma_{|AB'|}\,|AB'|}{\gamma_{|CB'|}\,|CB'|}\cdot\frac{\gamma_{|BC'|}\,|BC'|}{\gamma_{|AC'|}\,|AC'|}=\frac{\gamma_{a_1}a_1}{\gamma_{a_2}a_2}\cdot\frac{\gamma_{b_1}b_1}{\gamma_{b_2}b_2}\cdot\frac{\gamma_{c_1}c_1}{\gamma_{c_2}c_2}=1.\qquad(4)$$

From (1) and (4), we have

$$\frac{\gamma_{|AP|}\,|AP|}{\gamma_{|PA'|}\,|PA'|}=\frac{\gamma_a a}{2}\left(\frac{\gamma_{b_1}b_1\gamma_{c_2}c_2}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2}\right)+\frac{\gamma_a a}{2}\left(\frac{\gamma_{b_1}b_1\gamma_{c_2}c_2}{\gamma_{a_1}a_1\gamma_{b_1}b_1\gamma_{c_1}c_1}\right)=\frac{\gamma_a a}{2}\cdot\frac{2\gamma_{b_1}b_1\gamma_{c_2}c_2}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2}$$

$$=\frac{\gamma_a a\gamma_{b_1}b_1\gamma_{c_2}c_2}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2}.\qquad(5)$$

Similary we obtain that

$$\frac{\gamma_{|BP|}\,|BP|}{\gamma_{|PB'|}\,|PB'|}=\frac{\gamma_b b\gamma_{c_1}c_1\gamma_{a_2}a_2}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2},\qquad(6)$$

and

$$\frac{\gamma_{|CP|}\,|CP|}{\gamma_{|PC'|}\,|PC'|}=\frac{\gamma_c c\gamma_{a_1}a_1\gamma_{b_2}b_2}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2}.\qquad(7)$$

From the relations (5), (6) and (7) we get

$$P=\frac{\gamma_a a\gamma_{b_1}b_1\gamma_{c_2}c_2\cdot\gamma_b b\gamma_{c_1}c_1\gamma_{a_2}a_2\cdot\gamma_c c\gamma_{a_1}a_1\gamma_{b_2}b_2}{\left(\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2\right)^3}=\frac{\gamma_a a\gamma_b b\gamma_c c}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2}\qquad(8)$$

and

$$S=\frac{\gamma_a a\gamma_{b_1}b_1\gamma_{c_2}c_2+\gamma_b b\gamma_{c_1}c_1\gamma_{a_2}a_2+\gamma_c c\gamma_{a_1}a_1\gamma_{b_2}b_2}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2}.\qquad(9)$$

Because $\gamma_a\geqslant\gamma_{a_2},\gamma_b\geqslant\gamma_{b_2}$, and $\gamma_c\geqslant\gamma_{c_2}$ result

$$\gamma_a\gamma_b\gamma_c\geqslant\gamma_{a_2}\gamma_{b_2}\gamma_{c_2}.\qquad(10)$$

Therefore

$$\frac{\gamma_a a\gamma_b b\gamma_c c}{\gamma_{a_2}a_2\gamma_{b_2}b_2\gamma_{c_2}c_2}\geqslant 1.\qquad(11)$$

From the relations (8) and (11), we obtain that $P \geqslant 1$. If we use the inequality of arithmetic and geometric means, we obtain

$$S \geqslant 3\sqrt[3]{\frac{\gamma_a a\gamma_{b_1} b_1\gamma_{c_2} c_2 \cdot \gamma_b b\gamma_{c_1} c_1\gamma_{a_2} a_2 \cdot \gamma_c c\gamma_{a_1} a_1\gamma_{b_2} b_2}{\left(\gamma_{a_2} a_2\gamma_{b_2} b_2\gamma_{c_2} c_2\right)^3}} =$$

$$= 3\sqrt[3]{\frac{\gamma_a a\gamma_b b\gamma_c c}{\gamma_{a_2} a_2\gamma_{b_2} b_2\gamma_{c_2} c_2}}. \qquad (12)$$

From the relations (11) and (12), we obtain that $S \geqslant 3$. □

3 Conclusion

The special theory of relativity as was originally formulated by Einstein in 1905, [8], to explain the massive experimental evidence against ether as the medium for propagating electromagnetic waves, and Varičak in 1908 discovered the connection between special theory of relativity and hyperbolic geometry, [9]. The Einstein relativistic velocity model is another model of hyperbolic geometry. Many of the theorems of Euclidean geometry are relatively similar form in the Einstein relativistic velocity model, Smarandache minimum theorem is an example in this respect. In the Euclidean limit of large s, $s \to \infty$, gamma factor γ_v reduces to 1, so that the gyroinequalities (11) and (12) reduces to the

$$\frac{PA}{PA'} \cdot \frac{PB}{PB'} \cdot \frac{PC}{PC'} \geqslant 1,$$

and

$$\frac{PA}{PA'} + \frac{PB}{PB'} + \frac{PC}{PC'} \geqslant 3,$$

in Euclidean geometry. We observe that the previous inequalities are "weaker" than the inequalities of Smarandache's theorem of minimum.

References

1. Smarandache F. Nine Solved and Nine Open Problems in Elementary Geometry. arxiv.org/abs/1003.2153.
2. Ungar A. A. Analytic Hyperbolic Geometry and Albert Einstein's Special Theory of Relativity. Hackensack, NJ: World Scientific, 2008.
3. Ungar A. A. A Gyrovector Space Approach to Hyperbolic Geometry. Morgan & Claypool Publishers, 2009.
4. Barbu C., Pişcoran L. Van Aubel's Theorem in the Einstein Relativistic Velocity Model of Hyperbolic Geometry. Analele Universităţii din Timişoara, (to appear).
5. Ungar A. A. Analytic Hyperbolic Geometry Mathematical Foundations and Applications. World Scientific, Hackensack, 2005.
6. Barbu C. Fundamental Theorems of Triangle Geometry. Ed. Unique, Bacău, 2008 (in romanian).
7. Smarandache F. Problémes avec et sans... probléms! Somipress, Morocoo, 1983.
8. Einstein A. Relativity the special and general theory. Crown Publishers, New York, 1961.
9. Varičak V. Beitrage zur nichteuklidischen Geometrie. Jahresberichte des Deutschen Mathematiker Verein, 1908.

S-Denying a Theory

Florentin Smarandache
Department of Mathematics, University of New Mexico, Gallup, NM 87301, USA. E-mail: smarand@unm.edu

In this paper we introduce the operators of *validation* and *invalidation* of a proposition, and we extend the operator of *S-denying* a proposition, or an axiomatic system, from the geometric space to respectively any theory in any domain of knowledge, and show six examples in geometry, in mathematical analysis, and in topology.

1 Definitions

Let **T** be a theory in any domain of knowledge, endowed with an ensemble of sentences **E**, on a given space **M**.

E can be for example an axiomatic system of this theory, or a set of primary propositions of this theory, or all valid logical formulas of this theory, etc. **E** should be closed under the logical implications, i.e. given any subset of propositions $\mathbf{P_1}$, $\mathbf{P_2}$, ... in this theory, if **E** is a logical consequence of them then **Q** must also belong to this theory.

A sentence is a logic formula whose each variable is quantified {i.e. inside the scope of a quantifier such as: $\exists$ (*exist*), $\forall$ ($forall$), modal logic quantifiers, and other various modern logics' quantifiers}.

With respect to this theory, let **P** be a proposition, or a sentence, or an axiom, or a theorem, or a lemma, or a logical formula, or a statement, etc. of **E**.

It is said that **P** is *S-denied*[*] on the space **M** if **P** is valid for some elements of **M** and invalid for other elements of **M**, or **P** is only invalid on **M** but in at least two different ways.

An ensemble of sentences **E** is considered *S-denied* if at least one of its propositions is *S-denied*.

And a theory **T** is *S-denied* if its ensemble of sentences is *S-denied*, which is equivalent to at least one of its propositions being *S-denied*.

The proposition **P** is partially or totally denied/negated on **M**. The proposition **P** can be simultaneously validated in one way and invalidated in (finitely or infinitely) many different ways on the same space **M**, or only invalidated in (finitely or infinitely) many different ways.

The **invalidation** can be done in many different ways.

For example the statement **A** = "$x \neq 5$" can be invalidated as "$x = 5$" (total negation), but "$x \in \{5, 6\}$" (partial negation).

(Use a notation for *S-denying*, for invalidating in a way, for invalidating in another way a different notation; consider it as an operator: neutrosophic operator? A notation for invalidation as well.)

But the statement **B** = "$x > 3$" can be invalidated in many ways, such as "$x \leq 3$", or "$x = 3$", or "$x < 3$", or "x = -7", or "$x = 2$", etc. A negation is an invalidation, but not reciprocally – since an invalidation signifies a (partial or total) degree of negation, so invalidation may not necessarily be a complete negation. The negation of **B** is $\neg\mathbf{B}$ = "$x \leq 3$", while "x = -7" is a partial negation (therefore an invalidation) of **B**.

Also, the statement **C** = "John's car is blue and Steve's car is red" can be invalidated in many ways, as: "John's car is yellow and Steve's car is red", or "John's car is blue and Steve's car is black", or "John's car is white and Steve's car is orange", or "John's car is not blue and Steve's car is not red", or "John's car is not blue and Steve's car is red", etc.

Therefore, we can *S-deny* a theory in finitely or infinitely many ways, giving birth to many partially or totally denied versions/deviations/alternatives theories: $\mathbf{T_1}$, $\mathbf{T_2}$, ... These new theories represent **degrees of negations** of the original theory **T**.

Some of them could be useful in future development of sciences.

Why do we study such ***S-denying* operator**? Because our reality is heterogeneous, composed of a multitude of spaces, each space with different structures. Therefore, in one space a statement may be valid, in another space it may be invalid, and invalidation can be done in various ways. Or a proposition may be false in one space and true in another space or we may have a degree of truth and a degree of falsehood and a degree of indeterminacy. Yet, we live in this mosaic of distinct (even opposite structured) spaces put together.

S-denying involved the creation of the multi-space in geometry and of the *S-geometries* (1969).

It was spelt *multi-space*, or *multispace*, of *S-multispace*, or *mu-space*, and similarly for its: *multi-structure*, or *multistructure*, or *S-multistructure*, or *mu-structure*.

[*]The multispace operator S-denied (*Smarandachely-denied*) has been inherited from the previously published scientific literature (see for example Ref. [1] and [2]).

2 Notations

Let <A> be a statement (or proposition, axiom, theorem, etc.).

a) For the classical Boolean logic *negation* we use the same notation. The negation of <A> is noted by $\neg\mathbf{A}$ and $\neg\mathbf{A} = <nonA>$.

An *invalidation* of <A> is noted by **i(A)**, while a *validation* of <A> is noted by **v(A)**:

$$i(A) \subseteq 2^{<nonA>} \setminus \{\emptyset\} \text{ and } v(A) \subseteq 2^{<A>} \setminus \{\emptyset\}$$

where 2^X means the power-set of X, or all subsets of X.

All possible invalidations of <A> form a set of invalidations, notated by $I(A)$. Similarly for all possible validations of <A> that form a set of validations, and noted by $V(A)$.

b) *S-denying* of <A> is noted by $\mathbf{S\neg(A)}$. S-denying of <A> means some validations of <A> together with some invalidations of <A> in the same space, or only invalidations of <A> in the same space but in many ways.

Therefore, $S\neg(A) \subseteq V(A) \cup I(A)$ or $S\neg(A) \subseteq I(A)^k$, for $k \geq 2$.

3 Examples

Let's see some models of *S-denying*, three in a geometrical space, and other three in mathematical analysis (calculus) and topology.

3.1 The first *S-denying* model was constructed in 1969. This section is a compilation of ideas from paper [1].

An axiom is said *Smarandachely denied* if the axiom behaves in at least two different ways within the same space (i.e., validated and invalided, or only invalidated but in multiple distinct ways).

A *Smarandache Geometry* [SG] is a geometry which has at least one Smarandachely denied axiom.

Let's note any point, line, plane, space, triangle, etc. in such geometry by s-point, s-line, s-plane, s-space, s-triangle respectively in order to distinguish them from other geometries.

Why these hybrid geometries? Because in reality there do not exist isolated homogeneous spaces, but a mixture of them, interconnected, and each having a different structure.

These geometries are becoming very important now since they combine many spaces into one, because our world is not formed by perfect homogeneous spaces as in pure mathematics, but by non-homogeneous spaces. Also, SG introduce the degree of negation in geometry for the first time [for example an axiom is denied 40% and accepted 60% of the space] that's why they can become revolutionary in science and it thanks to the idea of partial denying/accepting of axioms/propositions in a space (making multi-spaces, i.e. a space formed by combination of many different other spaces), as in fuzzy logic the degree of truth (40% false and 60% true).

They are starting to have applications in physics and engineering because of dealing with non-homogeneous spaces.

The first model of S-denying and of SG was the following:

The axiom that through a point exterior to a given line there is only one parallel passing through it [Euclid's Fifth Postulate], was S-denied by having in the same space: no parallel, one parallel only, and many parallels.

In the Euclidean geometry, also called parabolic geometry, the fifth Euclidean postulate that there is only one parallel to a given line passing through an exterior point, is kept or validated.

In the Lobachevsky-Bolyai-Gauss geometry, called hyperbolic geometry, this fifth Euclidean postulate is invalidated in the following way: there are infinitely many lines parallels to a given line passing through an exterior point.

While in the Riemannian geometry, called elliptic geometry, the fifth Euclidean postulate is also invalidated as follows: there is no parallel to a given line passing through an exterior point.

Thus, as a particular case, Euclidean, Lobachevsky-Bolyai-Gauss, and Riemannian geometries may be united altogether, in the same space, by some SG's. These last geometries can be partially Euclidean and partially Non-Euclidean simultaneously.

3.2 Geometric Model (particular case of SG)

Suppose we have a rectangle $ABCD$.

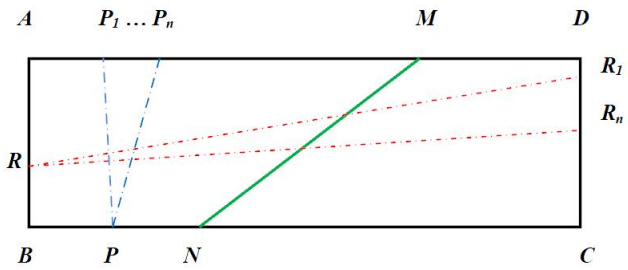


Fig. 1.

In this model we define as:

Point = any point inside or on the sides of this rectangle;

Line = a segment of line that connects two points of opposite sides of the rectangle;

Parallel lines = lines that do not have any common point (do not intersect);

Concurrent lines = lines that have a common point.

Let's take the line MN, where M lies on side AD and N on side BC as in the above *Fig. 1*. Let P be a point on side BC, and R a point on side AB.

Through P there are passing infinitely many parallels (PP_1, ..., PP_n, ...) to the line MN, but through R there is no parallel to the line MN (the lines RR_1, ..., RR_n cut line MN).

Therefore, the Fifth Postulate of Euclid (that though a point exterior to a line, in a given plane, there is only one parallel to that line) in *S-denied* on the space of the rectangle *ABCD* since it is invalidated in two distinct ways.

3.3 Another Geometric Model (another particular case of SG)

We change a little the Geometric Model 1 such that:

The rectangle *ABCD* is such that side *AB* is smaller than side *BC*. And we define as *line* the arc of circle inside (and on the borders) of *ABCD*, centered in the rectangle's vertices *A*, *B*, *C*, or *D*.

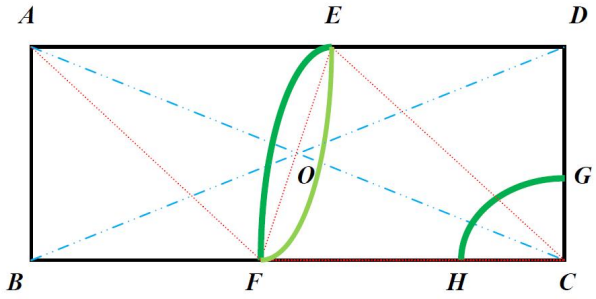


Fig. 2.

The axiom that: through two distinct points there exists only one line that passes through is S-denied (in three different ways):

a) Through the points *A* and *B* there is **no passing line** in this model, since there is no arc of circle centered in *A*, *B*, *C*, or *D* that passes through both points. See *Fig. 2*.

b) We construct the perpendicular $EF \perp AC$ that passes through the point of intersection of the diagonals *AC* and *BD*. Through the points *E* and *F* there are **two distinct lines** the dark green (left side) arc of circle centered in *C* since $CE \equiv FC$, and the light green (right side) arc of circle centered in *A* since $AE \equiv AF$, and because the right triangles $\sqcup\, COE$, $\sqcup\, COF$, $\sqcup\, AOE$, and $\sqcup\, AOF$ are all four congruent, we get $CE \equiv FC \equiv AE \equiv AF$.

c) Through the points *G* and *H* such that $CG \equiv CH$ (their lengths are equal) there is only **one passing line** (the dark green arc of circle *GH*, centered in *C*) since $AG \neq AH$ (their lengths are different), and similarly $BG \neq BH$ and $DG \neq DH$.

3.4 Example for the Axiom of Separation

The Axiom of Separation of Hausdorff is the following:

$$\forall\, x,y \in M\ \exists\, N(x), N(y)\colon\ N(x) \cap N(y) = \emptyset,$$

where *N(x)* is a neighborhood of *x*, and respectively *N(y)* is a neighborhood of *y*.

We can *S-deny* this axiom on a space *M* in the following way:

a) $\exists\, x_1, y_1 \in M$: $\exists\, N_1(x_1), N_1(y_1)$: $N_1(x_1) \cap N_1(y_1) = \emptyset$ where $N_1(x_1)$ is a neighborhood of x_1, and respectively $N_1(y_1)$ is a neighborhood of y_1; [validated].

b) $\exists\, x_2, y_2 \in M$: $\forall\, N_2(x_2), N_2(y_2)$: $N_2(x_2) \cap N_2(y_2) \neq \emptyset$; where $N_2(x_2)$ is a neighborhood of x_2, and respectively $N_2(y_2)$ is a neighborhood of y_2; [invalidated].

Therefore we have two categories of points in *M*: some points that verify The Axiom of Separation of Hausdorff and other points that do not verify it. So *M* becomes a partially separable and partially inseparable space, or we can see that *M* has some **degrees of separation**.

3.5 Example for the Norm

If we remove one or more axioms (or properties) from the definition of a notion <*A*> we get a pseudo-notion <*pseudoA*>.

For example, if we remove the third axiom (inequality of the triangle) from the definition of the <*norm*> we get a <*pseudonorm*>.

The axioms of a **norm** on a real or complex vectorial space *V* over a field *F*, $x \mapsto \|.\|$, are the following:

a) $\|x\| = 0 \Leftrightarrow x = 0$.

b) $\forall\, x \in V, \forall\, \alpha \in F, \|\alpha \cdot x\| = |\alpha| \cdot \|x\|$.

c) $\forall\ x,\ y \in V,\ \|x + y\| \leq \|x\| \cdot \|y\|$ (inequality of the triangle).

For example, a **pseudo-norm** on a real or complex vectorial space *V* over a field *F*, $x \mapsto_p \|.\|$, may verify only the first two above axioms of the norm.

A pseudo-norm is a particular case of an *S-denied norm* since we may have vectorial spaces over some given scalar fields where there are some vectors and scalars that satisfy the third axiom [validation], but others that do not satisfy [invalidation]; or for all vectors and scalars we may have either $\|x + y\| = 5 \cdot \|x\| \cdot \|y\|$ or $\|x + y\| = 6 \cdot \|x\| \cdot \|y\|$, so invalidation (since we get $\|x + y\| > \|x\| \cdot \|y\|$) in two different ways.

Let's consider the complex vectorial space

$$\mathscr{C} = \{\ a+b \cdot i, \text{ where } a, b \in \mathbb{R},\ i = \sqrt{-1}\ \}$$

over the field of real numbers $\mathbb{R}$.

If $z = a+b \cdot i \in \mathscr{C}$ then its pseudo-norm is $\| z \| = \sqrt{a^2 + b^2}$. This verifies the first two axioms of the norm, but does not satisfy the third axiom of the norm since:

For $x = 0 + b \cdot i$ and $y = a + 0 \cdot i$ we get:

$$\|x + y\| = \|a + b \cdot i\| = \sqrt{a^2 + b^2} \leq \|x\| \cdot \|y\|$$
$$= \|0 + b \cdot i\| \cdot \|a + 0 \cdot i\| = |a \cdot b|, \text{ or } a^2 + b^2 \leq a^2 \cdot b^2.$$

But this is true for example when $a = b \geq \sqrt{2}$ (validation), and

false if one of a or b is zero and the other is strictly positive (invalidation).
Pseudo-norms are already in use in today's scientific research, because for some applications the norms are considered too restrictive.

Similarly one can define a pseudo-manifold (relaxing some properties of the manifold), etc.

3.6 Example in Topology

A topology $\mathcal{O}$ on a given set E is the ensemble of all parts of E verifying the following properties:

a) E and the empty set $\emptyset$ belong to $\mathcal{O}$.
b) Intersection of any two elements of $\mathcal{O}$ belongs to $\mathcal{O}$ too.
c) Union of any family of elements of $\mathcal{O}$ belongs to $\mathcal{O}$ too.

Let's go backwards. Suppose we have a topology $\mathcal{O}_1$ on a given set E_1, and the second or third (or both) previous axioms have been *S-denied*, resulting an ***S-denied topology*** $S\neg(\mathcal{O}_1)$ on the given set E_1.

In general, we can go back and "recover" (reconstruct) the original topology $\mathcal{O}_1$ from $S\neg(\mathcal{O}_1)$ by recurrence: if two elements belong to $S\neg(\mathcal{O}_1)$ then we set these elements and their intersection to belong to $\mathcal{O}_1$, and if a family of elements belong to $S\neg(\mathcal{O}_1)$ then we set these family elements and their union to belong to $\mathcal{O}_1$; and so on: we continue this recurrent process until it does not bring any new element to $\mathcal{O}_1$.

Conclusion

Decidability changes in an *S-denied theory*, i.e. a defined sentence in an S-denied theory can be partially deducible and partially undeducible (we talk about **degrees of deducibility of a sentence** in an *S-denied theory*).

Since in classical deducible research, a theory T of language L is said complete if any sentence of L is decidable in T, we can say that an *S-denied theory* is partially complete (or has some **degrees of completeness** and degrees of incompleteness).

Submitted on September 22, 2010 / Accepted on September 26, 2010

References

1. Kuciuk L., Antholy M. An Introduction to Smarandache Geometries. Mathematics Magazine, Aurora, Canada, v. 12, 2003.
2. Mao, Linfan, An introduction to Smarandache geometries on maps. 2005 International Conference on Graph Theory and Combinatorics, Zhejiang Normal University, Jinhua, Zhejiang, P. R. China, June 25-30, 2005.
3. Rabounski D. Smarandache Spaces as a New Extension of the Basic Space-Time of General Relativity. *Progress in Physics*, 2010, v. 2, L1-L2.
4. Smarandache F. Paradoxist Geometry. State Archives from Valcea, Rm. Valcea, Romania, 1969; and in "Paradoxist Mathematics", Collected Papers (Vol. II), Kishinev University Press, Kishinev, 1997, pp. 5-28.
5. Smarandache F. Multi-Space and Multi-Structure, in: "Neutrosophy. Neutrosophic Logic, Set, Probability and Statistics", American Research Press, 1998.
6. Szudzik M., "Sentence", from *MathWorld* – A Wolfram Web Resource, created by Eric W. Weisstein, http://mathworld.wolfram.com/Sentence.html.

On the Quantum Mechanical State of the Δ^{++} Baryon

Eliahu Comay

Charactell Ltd. P.O. Box 39019, Tel Aviv 61390 Israel, E-mail: elicomay@post.tau.ac.il

The Δ^{++} and the Ω^{-} baryons have been used as the original reason for the construction of the Quantum Chromodynamics theory of Strong Interactions. The present analysis relies on the multiconfiguration structure of states which are made of several Dirac particles. It is shown that this property, together with the very strong spin-dependent interactions of quarks provide an acceptable explanation for the states of these baryons and remove the classical reason for the invention of color within Quantum Chromodynamics. This explanation is supported by several examples that show a Quantum Chromodynamics' inconsistency with experimental results. The same arguments provide an explanation for the problem called the proton spin crisis.

1 Introduction

It is well known that writing an atomic state as a sum of terms, each of which belongs to a specific configuration is a useful tool for calculating electronic properties of the system. This issue has already been recognized in the early days of quantum mechanics [1]. For this purpose, mathematical tools (called angular momentum algebra) have been developed mainly by Wigner and Racah [2]. Actual calculations have been carried out since the early days of electronic computers [3]. Many specific properties of atomic states have been proven by these calculations. In particular, these calculations have replaced guesses and conjectures concerning the mathematical form of atomic states by evidence based on a solid mathematical basis. In this work, a special emphasis is given to the following issue: Contrary to a naive expectation, even the ground state of a simple atom is written as a sum of more than one configuration. Thus, the apparently quite simple closed shell ground state of the two electron He atom, having $J^{\pi} = 0^{+}$, *disagrees* with the naive expectation where the two electrons are just in the $1s^2$ state. Indeed, other configurations where individual electrons take higher angular momentum states (like $1p^2$, $1d^2$, etc.) have a non-negligible part of the state's description [3]. The multiconfiguration description of the ground state of the He atom proves that shell model notation of state is far from being complete. Notation of this model can be regarded as an anchor configuration defining the J^{π} of the state. Therefore, all relevant configurations must have the same parity and their single-particle angular momentum must be coupled to the same J.

This paper discusses some significant elements of this scientific knowledge and explains its crucial role in a quantum mechanical description of the states of the Δ^{++}, the Δ^{-} and the Ω^{-} baryons. In particular, it is proved that these baryons do not require the introduction of new structures (like the $SU(3)$ color) into quantum mechanics. A by-product of this analysis is the settlement of the "proton spin crisis" problem.

The paper is organized as follows. The second section describes briefly some properties of angular momentum algebra. It is proved in the third section that ordinary laws of quantum mechanics explain why the states of the Δ^{++}, Δ^{-} and Ω^{-} baryons are consistent with the Pauli exclusion principle. This outcome is used in the fourth section for showing that QCD does not provide the right solution for hadronic states. The problem called "proton spin crisis" is explained in the fifth section. The last section contains concluding remarks.

2 Some Features of Angular Momentum Algebra

Consider the problem of a bound state of N Dirac particles. (Baryonic states are extremely relativistic. Therefore, a relativistic formulation is adopted from the beginning.) This system is described as an eigenfunction of the Hamiltonian. Thus, the time variable is removed and one obtains a problem of $3N$ spatial variables for each of the four components of a Dirac wave function. It is shown here how angular momentum algebra can be used for obtaining a dramatic simplification of the problem.

The required solution is constructed as a sum of terms, each of which depends on all the independent variables mentioned above. Now angular momentum is a good quantum number of a closed system and parity is a good quantum number for systems whose state is determined by strong or electromagnetic interactions. Thus, one takes advantage of this fact and uses only terms that have the required angular momentum and parity, denoted by J^{π}. (Later, the lower case j^{π} denotes properties of a bound spin-1/2 single particle.)

The next step is to write each term as a product of N single particle Dirac functions, each of which has a well defined angular momentum and parity. The upper and lower parts of a Dirac function have opposite parity [4, see p. 53]. The angular coordinates of the two upper components of the Dirac function have angular momentum l and they are coupled with the spin to an overall angular momentum $j = l \pm 1/2$ ($j > 0$). The two lower components have angular momentum $(l \pm 1) \geq 0$ and together with the spin, they are coupled to the same j. The spatial angular momentum eigenfunctions having an eignevalue l, make a set of $(2l + 1)$ members de-

noted by $Y_{lm}(\theta, \phi)$, where θ, ϕ denote the angular coordinates and $-l \le m \le l$ [2].

It is shown below how configurations can be used for describing a required state whose parity and overall spin are known.

3 The Δ^{++} State

The purpose of this section is to describe how the state of each of the four Δ baryons can be constructed in a way that abides by ordinary quantum mechanics of a system of three fermions. Since the $\Delta^{++}(1232)$ baryon has 3 valence quarks of the u flavor, the isospin $I = 3/2$ of all four Δ baryons is fully symmetric. Therefore, the space-spin components of the 3-particle terms must be antisymmetric. (An example of relevant nuclear states is presented at the end of this section.) Obviously, each of the 3-particle terms must have a total spin $J = 3/2$ and an even parity. For writing down wave functions of this kind, single particle wave functions having a definite j^π and appropriate radial functions are used. A product of n specific j^π functions is called a configuration and the total wave function takes the form of a sum of terms, each of which is associated with a configuration. Here n=3 and only even parity configurations are used. Angular momentum algebra is applied to the single particle wave functions and yields an overall $J = 3/2$ state. In each configuration, every pair of the Δ^{++} u quarks must be coupled to an antisymmetric state. r_j denotes the radial coordinate of the jth quark.

Each of the A-D cases described below contains one configuration and one or several antisymmetric 3-particle terms. The radial functions of these examples are adapted to each case.

Notation: $f_i(r_j)$, $g_i(r_j)$, $h_i(r_j)$ and $v_i(r_j)$ denote radial functions of Dirac single particle $1/2^+$, $1/2^-$, $3/2^-$ and $3/2^+$ states, respectively. The index i denotes the excitation level of these functions. Each of these radial functions is a two-component function, one for the upper 2-component spin and the other for the lower 2-component spin that belong to a 4-component Dirac spinor.

A. In the first example all three particles have the same j^π,

$$f_0(r_0)f_1(r_1)f_2(r_2)\ 1/2^+\ 1/2^+\ 1/2^+. \qquad (1)$$

Here the spin part is fully symmetric and yields a total spin of 3/2. The spatial state is fully antisymmetric. It is obtained from the 6 permutations of the three orthogonal $f_i(r_j)$ functions divided by $\sqrt{6}$. This configuration is regarded as the anchor configuration of the state.

B. A different combination of $j_i = 1/2$ can be used,

$$f_0(r_0)g_0(r_1)g_1(r_2)\ 1/2^+\ 1/2^-\ 1/2^-. \qquad (2)$$

Here, the two single particle $1/2^-$ spin states are coupled symmetrically to j=1 and they have two orthogonal radial functions g_i. The full expression can be antisymmetrized.

C. In this example, just one single particle is in an angular excitation state,

$$f_0(r_0)f_0(r_1)v_0(r_2)\ 1/2^+\ 1/2^+\ 3/2^+. \qquad (3)$$

Here we have two $1/2^+$ single particle functions having the same non-excited radial function. These spins are coupled antisymmetrically to a spin zero two particle state. These spins have the same non-excited radial function. The third particle yields the total $J = 3/2$ state. The full expression can be antisymmetrized.

D. Here all the three single particle j^π states take different values. Therefore, the radial functions are free to take the lowest level,

$$f_0(r_0)g_0(r_1)h_0(r_2)\ 1/2^+\ 1/2^-\ 3/2^-. \qquad (4)$$

Due to the different single particle spins, the antisymmetrization task of the spin coordinates can easily be done. (The spins can be coupled to a total $J = 3/2$ state in two different ways. Hence, two different terms belong to this configuration.)

The examples A-D show how a Hilbert space basis for the $J^\pi = 3/2^+$ state can be constructed for three identical fermions. Obviously, more configurations can be added and the problem can be solved by ordinary spectroscopic methods. It should be noted that unlike atomic states, the very strong spin dependent interactions of hadrons are expected to yield a higher configuration mixture.

It is interesting to note that a similar situation is found in nuclear physics. Like the u,d quarks, the proton and the neutron are spin-1/2 fermions belonging to an isospin doublet. This is the basis for the common symmetry of isospin states described below. Table 1 shows energy levels of each of the four A=31 nuclei examined [5, see p. 357]. Each of these nuclei has 14 protons and 14 neutrons that occupy a set of inner closed shells and three nucleons outside these shells. (The closed shells are $1/2^+$, $3/2^-$, $1/2^-$, and $5/2^+$. The next shells are the $1/2^+$ that can take 2 nucleons of each type and the

Table 1: Nuclear A=31 Isospin State Energy levels (MeV)

J^π	I (T)[a]	^{31}Si	^{31}P	^{31}S	^{31}Cl[b]
$1/2^+$	1/2	—	0	0	—
$3/2^+$	3/2	0	6.38	6.27	0
$1/2^+$	3/2	0.75	7.14	7.00	—

[a] I,T denote isospin in particle physics and nuclear physics, respectively.
[b] The ^{31}Cl data is taken from [6].

$3/2^+$ shell that is higher than the $1/2^+$ shell. [7, See p. 245]. The state is characterized by these three nucleons that define the values of total spin, parity and isospin. The first line of table 1 contains data of the ground states of the $I = 1/2$ ^{31}P and ^{31}S nuclei. The second line contains data of the lowest level of the $I = 3/2$ state of the four nuclei. The quite small energy difference between the ^{31}P and ^{31}S excited states illustrates the quite good accuracy of the isospin approximation. The third line of the table shows the first excited $I = 3/2$ state of each of the four nuclei. The gap between states of the third and the second lines is nearly, but not precisely, the same. It provides another example of the relative goodness of the isospin approximation.

The nuclear states described in the first and the second lines of table 1 are relevant to the nucleons and the Δ baryons of particle physics. Indeed, the states of both systems are characterized by three fermions that may belong to two different kinds and where isospin is a useful quantum number. Here the neutron and the proton correspond to the ground state of ^{31}P and ^{31}S, respectively, whereas energy states of the second line of the table correspond to the four Δ baryons. Every nuclear energy state of table 1 has a corresponding baryon that has the same spin, parity, isospin and the I_z isospin component. Obviously, the dynamics of the nuclear energy levels is completely different from hadronic dynamics. Indeed, the nucleons are composite spin-1/2 particles whose state is determined by the strong nuclear force. This is a residual force characterized by a rapidly decaying attractive force at the vicinity of the nucleon size and a strong repulsive force at a smaller distance [7, see p. 97]. On the other hand, the baryonic quarks are elementary pointlike spin-1/2 particles whose dynamics differs completely from that of the strong nuclear force. However, both systems are made of fermions and the spin, parity and isospin analogy demonstrates that *the two systems have the same internal symmetry.*

The following property of the second line of table 1 is interesting and important. Thus, all nuclear states of this line have the same *symmetric* $I = 3/2$ state. Hence, due to the Pauli exclusion principle, all of them have the same *antisymmetric space-spin state.* Now, for the ^{31}P and ^{31}S nuclei, this state is an excited state because they have lower states having $I = 1/2$. However, the ^{31}Si and ^{31}Cl nuclei have no $I = 1/2$ state, because their $I_z = 3/2$. Hence, the second line of table 1 shows *the ground state of the* $I_z = 3/2$ *nuclei*. It will be shown later that this conclusion is crucial for having a good understanding of an analogous quark system. Therefore it is called *Conclusion A*.

Now, the ^{31}Si has three neutrons outside the 28 nucleon closed shells and the ^{31}Cl has three protons outside these shells. Hence, the outer shell of these two nuclear states consists of three identical fermions which make the required ground state. Relying on this nuclear physics example, one deduces that *the Pauli exclusion principle is completely consistent with three identical fermions in a* $J^\pi = 3/2^+$ *and* $I = 3/2$ *ground state.* The data of table 1 are well known in nuclear physics.

A last remark should be made before the end of this section. As explained in the next section, everything said above on the isospin quartet $J^\pi = 3/2^+$ states of the three *ud* quark flavor that make the four Δ baryons, holds for the full decuplet of the three *uds* quarks, where, for example, the Ω^- state is determined by the three *sss* quarks. In particular, like the Δ^{++} and the Δ^-, *the* Ω^- *baryon is the ground state of the three sss quarks and each of the baryons of the decuplet has an antisymmetric space-spin wave function.*

4 Discussion

It is mentioned above that spin-dependent interactions are much stronger in hadronic states than in electronic states. This point is illustrated by a comparison of the singlet and triplet states of the positronium [8] with the π^0 and ρ^0 mesons [9]. The data are given in table 2. The fourth column of the table shows energy difference between each of the $J^\pi = 1^-$ states and the corresponding $J^\pi = 0^-$ state. The last column shows the ratio between this difference and the energy of the $J^\pi = 0^-$ state.

Both electrons and quarks are spin-1/2 pointlike particles. The values of the last column demonstrate a clear difference between electrically charged electrons and quarks that participate in strong interactions. Indeed, the split between the two electronic states is *very small*. This is the reason for the notation of *fine structure* for the spin dependent split between electronic states of the same excitation level [10, see p. 225]. Table 2 shows that the corresponding situation in quark systems is larger by more than 9 orders of magnitude. Hence, spin dependent interactions in hadrons are very strong and make an important contribution to the state's energy.

Now, electronic systems in atoms satisfy the Hund's rules [10, see p. 226]. This rule says that in a configuration, the state having the highest spin is bound stronger. Using this rule *and* the very strong spin-dependent hadronic interaction which is demonstrated in the last column of table 2, one con-

Table 2: Positronium and meson energy (MeV)

Particle	J^π	Mass	$M(1^-) - M(0^-)$	$\Delta M/M(0^-)$
e^+e^-	0^-	~ 1.022	—	—
e^+e^-	1^-	~ 1.022	8.4×10^{-10}	8.2×10^{-10}
π^0	0^-	135	—	—
ρ^0	1^-	775	640	4.7

cludes that the anchor configuration A of the previous section really describes a very strongly bound state of the Δ^{++} baryon. In particular, the isospin doublet $J^\pi = 1/2^+$ state of the neutron and the proton correspond to the same $J^\pi = 1/2^+$ of the ground state of the $A = 31$ nuclei displayed in the first line of table 1. The isospin quartet of the Δ baryons correspond to the isospin quartet of the four nuclear states displayed in the second line of table 1.

Here the significance of Conclusion A of the previous section becomes clear. Indeed, an analogy is found between the two nuclear states of the $I = 3/2$ and $I_z = \pm 1/2$, namely the ^{31}P and the ^{31}S are *excited states* of these nuclei whereas the $I = 3/2$ and $I_z = \pm 3/2$, namely the ^{31}Si and the ^{31}Cl states are the *ground states* of these nuclei. The same pattern is found in the particle physics analogue. The Δ^+ and the Δ^0 are *excited states* of the proton and the neutron, respectively. This statement relies on the fact that both the proton and the Δ^+ states are determined by the *uud* quarks. Similarly, the neutron and the Δ^0 states are determined by the *udd* quarks. On the other hand, in the case of the ^{31}Si and the ^{31}Cl nuclei, the $I = 3/2$ and $J^\pi = 3/2^+$ states are the *ground states* of these nuclei. The same property holds for the Δ^{++} and the Δ^-, which are the *ground states* of the *uuu* and *ddd* quark systems, respectively.

The relationship between members of the lowest energy $J^\pi = 1/2^+$ baryonic octet and members of the $J^\pi = 3/2^+$ baryonic decuplet can be described as follows. There are 7 members of the decuplet that are excited states of corresponding members of the octet. Members of each pair are made of the same specific combination of the *uds* quarks. The Δ^{++}, Δ^- and Ω^- baryons have no counterpart in the octet. Thus, in spite of being a part of the decuplet whose members have space-spin antisymmetric states, these three baryons are the *ground state* of the *uuu, ddd* and *sss* quarks, respectively.

This discussion can be concluded by the following statements: *The well known laws of quantum mechanics of identical fermions provide an interpretation of the* Δ^{++}, Δ^- *and* Ω^- *baryons, whose state is characterized by three uuu, ddd and sss quarks, respectively. There is no need for any fundamental change in physics in general and for the invention of color in particular. Like all members of the decuplet, the states of these baryons abide by the Pauli exclusion principle.* Hence, one wonders why particle physics textbooks regard precisely the same situation of the four Δ baryons as a fiasco of the Fermi-Dirac statistics [11, see p. 5].

The historic reasons for the QCD creation are the states of the Δ^{++} and the Ω^- baryons. These baryons, each of which has three quarks of the same flavor, are regarded as the *classical reason* for the QCD invention [12, see p. 338]. The analysis presented above shows that this argument does not hold water. For this reason, one wonders whether QCD is really a correct theory. This point is supported by the following examples which show that QCD is inconsistent with experimental results.

1. The interaction of hard real photons with a proton is practically the same as its interaction with a neutron [13]. This effect cannot be explained by the photon interaction with the nucleons' charge constituents, because these constituents take different values for the proton and the neutron. The attempt to recruit Vector Meson Dominance (VMD) for providing an explanation is unacceptable. Indeed, Wigner's analysis of the irreducible representations of the Poincare group [14] and [15, pp. 44–53] proves that VMD, which mixes a massive meson with a massless photon, is incompatible with Special relativity. Other reasons of this kind also have been published [16].
2. QCD experts have predicted the existence of strongly bound pentaquarks [17, 18]. In spite of a long search, the existence of pentaquarks has not been confirmed [19]. By contrast, correct physical ideas used for predicting genuine particles, like the positron and the Ω^-, have been validated by experiment after very few years (and with a technology which is very very poor with respect to that used in contemporary facilities).
3. QCD experts have predicted the existence of chunks of Strange Quark Matter (SQM) [20]. In spite of a long search, the existence of SQM has not been confirmed [21].
4. QCD experts have predicted the existence of glueballs [22]. In spite of a long search, the existence of glueballs has not been confirmed [9].
5. For a very high energy, the proton-proton (*pp*) total and elastic cross section increase with collision energy [9] and the ratio of the elastic cross section to the total cross section is nearly a constant which equals about 1/6. This relationship between the *pp* cross sections is completely different from the high energy electron-proton (*ep*) scattering data where the total cross section decreases for an increasing collision energy and the elastic cross section's portion becomes negligible [23]. This effect proves that the proton contains a quite solid component that can take the heavy blow of the high energy collision and keep the entire proton intact. This object cannot be a quark, because in energetic *ep* scattering, the electron strikes a single quark and the relative part of elastic events is negligible. QCD has no explanation for the *pp* cross section data [24].

5 The Proton Spin Crisis

Another problem which is settled by the physical evidence described above is called *the proton spin crisis* [25, 26]. Here polarized muons have been scattered by polarized protons. The results prove that the instantaneous quark spin sums up to a very small portion of the entire proton's spin. This outcome, which has been regarded as a surprise [26], was later

supported by other experiments. The following lines contain a straightforward explanation of this phenomenon.

The four configurations A-D that are a part of the Hilbert space of the Δ^{++} baryon are used as an illustration of the problem. Thus, in configuration A, all single particle spins are parallel to the overall spin. The situation in configuration B is different. Here the single particle function $j^{\pi} = 1/2^{-}$ is a coupling of $l = 1$ and $s = 1/2$. This single particle coupling has two terms whose numerical values are called Clebsh-Gordan coefficients [2]. In one of the coupling terms, the spin is parallel to the overall single particle angular momentum and in the other term it is antiparallel to it. This is an example of an internal partial cancellation of the contribution of the single particle spin to the overall angular momentum. (As a matter of fact, this argument also holds for the lower pair of components of each of Dirac spinor of configuration A. Here the lower pair of the high relativistic system is quite large and it is made of $l = 1$ $s = 1/2$ coupled to $J = 1/2$.) In configuration C one finds the same effect. Spins of the first and the second particles are coupled to $j_{01} = 0$ and cancel each other. In the third particle the $l = 2$ spatial angular momentum is coupled with the spin to $j = 3/2$. Here one also finds two terms and the contribution of their single-particle spin-1/2 partially cancels. The same conclusion is obtained from an analogous analysis of the spins of configuration D.

It should be pointed out that the very large hadronic spin-dependent interaction which is demonstrated by the data of table 2, proves that in hadronic states, one needs many configurations in order to construct a useful basis for the Hilbert space of a baryon. It follows that in hadrons the internal single particle cancellation seen in configurations of the previous section, is expected to be quite large.

Obviously, the configuration structure of the proton relies on the same principles. Here too, many configurations, each of which has the quantum numbers $J^{\pi} = 1/2^{+}$, are needed for the Hilbert space. Thus, a large single particle spin cancellation is obtained. Therefore, the result of [25] is obvious. It should make neither a crisis nor a surprise.

On top of what is said above, the following argument indicates that the situation is more complicated and the number of meaningful configurations is even larger. Indeed, it has been shown that beside the three valence quarks, the proton contains additional quark-antiquark pair(s) whose probability is about 1/2 pair [23, see p. 282]. It is very reasonable to assume that all baryons share this property. The additional quark-antiquark pair(s) increase the number of useful configurations and of their effect on the Δ^{++} ground state and on the proton spin as well.

6 Concluding Remarks

Relying on the analysis of the apparently quite simple ground state of the He atomic structure [3], it is argued here that *many* configurations are needed for describing a quantum mechanical state of more than one Dirac particle. This effect is much stronger in baryons. where, as shown in table 2, spin-dependent strong interactions are very strong indeed. This effect and the multiconfiguration basis of hadronic states do explain the isospin quartet of the $J = 3/2^{+}$ Δ baryons. Here the Δ^0 and the Δ^+ are *excited states* of the neutron and the proton, respectively whereas their isospin counterparts, the Δ^{++} and the Δ^- are *ground states* of the *uuu* and the *ddd* quark systems, respectively. Analogous conclusions hold for all members of the $J = 3/2^{+}$ baryonic decuplet that includes the Ω^- baryon. It is also shown that states of four $A = 31$ nuclei are analogous to the nucleons and the Δs isospin quartet (see table 1).

The discussion presented above shows that there is no need for introducing a new degree of freedom (like color) in order to settle the states of Δ^{++}, Δ^- and Ω^- baryons with the Pauli exclusion principle. Hence, there is no reason for the QCD invention. Several inconsistencies of QCD with experimental data support this conclusion.

Another aspect of recognizing implications of the multi-configuration structure of hadrons is that the proton spin crisis experiment, which shows that instantaneous spins of quarks make a little contribution to the proton's spin [25], creates neither a surprise nor a crisis.

Submitted on November 7, 2010 / Accepted on November 9, 2010

References

1. Wong S. S. M. Nuclear Physics. Wiley, New York, 1998.
2. de-Shalit A. and Talmi I. Nuclear Shell Theory. Academic, New York, 1963.
3. Weiss A. W. Configuration Interaction in Simple Atomic Systems. *Physcal Review*, 1961, v. 122, 1826–1836.
4. Bjorken J. D. and Drell S. D. Relativistic Quantum Mechanics. McGraw, New York, 1964.
5. Endt P. M. Energy Levels of A = 21–44 Nuclei (VII). *Nuclear Physics A*, 1990, v. 521, 1–830.
6. Kankainen A. et al. Excited states in ^{31}S studied via beta decay of ^{31}Cl. *The European Physical Journal A — Hadrons and Nuclei*, 2006, v. 27, 67–75.
7. Wong S. S. M. Nuclear Physics. Wiley, New York, 1998.
8. Berko S. and Pendleton H. N. Positronium. *Annual Review of Nuclear and Particle Science*, 1980, v. 30, 543–581.
9. Nakamura K. et al. (Particle Data Group), Review of Particle Physics 075021. *Journal of Physics G: Nuclear and Particle Physics*, 2010, v. 37, 1–1422.
10. Landau L. D. and Lifshitz E. M. Quantum Mechanics. Pergamon, London, 1959.
11. Halzen F. and Martin A. D. Quarks and Leptons. Wiley, New York, 1984.
12. Close F. E. An Introduction to quarks and Partons. Academic, London, 1979.
13. Bauer T. H., Spital R. D., Yennie D. R. and Pipkin F. M. The Hadronic Properties of the Photon in High-Energy Interactions. *Reviews of Modern Physics*, 1978, v. 50, 261–436.
14. Wigner E. P. On Unitary Representations of the Inhomogeneous Lorentz Group. *The Annals of Mathematics*, 1939, v. 40, 149–204.

15. Schweber S. S. An Introduction to Relativistic Quantum Field Theory. Harper & Row, New York, 1964.
16. Comay E. Remarks on Photon-Hadron Interactions. *Apeiron*, 2003, v. 10 (2), 87–103.
17. Gignoux C., Silvestre-Brac B. and Richard J. M. Possibility of Stable Multiquark Baryons. *Physics Letters B*, 1987, v. 193, 323–326.
18. Lipkin H. J. New Possibilities for Exotic Hadrons–Anticharmed Strange Baryons. *Physics Letters B*, 1987, v. 195, 484–488.
19. Whol C. G. In the 2009 report of PDG at http://pdg.lbl.gov/2009/reviews/rpp2009-rev-pentaquarks.pdf
20. Witten E. Cosmic Separation of Phases. *Physical Review D*, 1984, v. 30, 272–285.
21. Han K. et al. Search for Stable Strange Quark Matter in Lunar Soil. *Physical Review Letters*, 2009, v. 103, 092302-1–092302-4.
22. Frauenfelder H. and Henley E. M. Subatomic Physics. Prentice Hall, Englewood Cliffs, 1991.
23. Perkins D. H. Introductions to High Energy Physics. Addison-Wesley, Menlo Park, 1987.
24. Arkhipov A. A. On Global Structure of Hadronic Total Cross Sections. arXiv: hep-ph/9911533v2
25. Ashman J. et al. (EMC) A Measurement of the Spin Asymmetry and Determination of the Structure Function g_1 in Deep Inelastic Muon-Proton Scattering. *Physics Letters B*, 1988, v. 206, 364–370.
26. Myhrer F. and Thomas A. W. Understanding the Proton's Spin Structure. *Journal of Physics G*, 2010, v. 37, 023101-1–023101–12.

Sound-Particles and Phonons with Spin 1

Vahan Minasyan and Valentin Samoilov
Scientific Center of Applied Research, JINR, Dubna, 141980, Russia
E-mails: mvahan@scar.jinr.ru; scar@off-serv.jinr.ru

We present a new model for solids which is based on the stimulated vibration of independent neutral Fermi-atoms, representing independent harmonic oscillators with natural frequencies, which are excited by actions of the longitudinal and transverse elastic waves. Due to application of the principle of elastic wave-particle duality, we predict that the lattice of a solid consists of two type Sound Boson-Particles with spin 1 with finite masses. Namely, these lattice Boson-Particles excite the longitudinal and transverse phonons with spin 1. In this letter, we estimate the masses of Sound Boson-Particles which are around 500 times smaller than the atom mass.

1 Introduction

The original theory proposed by Einstein in 1907 was of great historical relevance [1]. In the Einstein model, each atom oscillates relatively to its neighbors in the lattice which execute harmonic motions around fixed positions, the knots of the lattice. He treated the thermal property of the vibration of a lattice of N atoms as a $3N$ harmonic independent oscillator by identical own frequency Ω_0 which was quantized by application of the prescription developed by Plank in connection with the theory of Black Body radiation. The Einstein model could obtain the Dulong and Petit prediction at high temperature but could not reproduce an adequate representation of the the lattice at low temperatures. In 1912, Debye proposed to consider the model of the solid [2], by suggestion that the frequencies of the $3N$ harmonic independent oscillators are not equal as it was suggested by the Einstein model. In addition to his suggestion, the acoustic spectrum of solid may be treated as if the solid represented a homogeneous medium, except that the total number of independent elastic waves is cut off at $3N$, to agree with the number of degrees of freedom of N atoms. In this respect, Debye stated that one longitudinal and two transverse waves are excited in solid. These velocities of sound cannot be observed in a solid at frequencies above the cut-off frequency. Also, he suggested that phonon is a spinless. Thus, the Debye model correctly showed that the heat capacity is proportional to the T^3 law at low temperatures. At high temperatures, he obtained the Dulong-Petit prediction compatible to experimental results.

In this letter, we propose a new model for solids which consists of neutral Fermi-atoms, fixed in the knots of lattice. In turn, within the formalism of Debye, we may predict that lattice represents as the Bose-gas of Sound-Particles with finite masses m_l and m_t, corresponding to a longitudinal and a transverse elastic field. In this sense, the lattice is considered as a new substance of matter consisting of Sound-Particles, which excite the one longitudinal and one transverse elastic waves (this approach is differ from Debye one). These waves act on the Fermi-atoms which are vibrating with the natural frequencies Ω_l and Ω_t. Thus, there are stimulated vibrations of the Fermi-atoms by under action of longitudinal and transverse phonons with spin 1. In this context, we introduce a new principle of elastic wave-particle duality, which allows us to build the lattice model. The given model leads to the same results as presented by Debye's theory.

2 Analysis

As we suggest, the transfer of heat from one part of the body to another occurs through the lattice. This process is very slow. Therefore, we can regard any part of the body as thermally insulated, and there occur adiabatic deformations. In this respect, the equation of motion for an elastic continuum medium [3] represents as

$$\varrho\ddot{\vec{u}}(\vec{r},t) = c_t^2\nabla^2\vec{u}(\vec{r},t) + (c_l^2 - c_t^2)\, grad \cdot div\, \vec{u}(\vec{r},t) \qquad (1)$$

where $\vec{u} = \vec{u}(\vec{r},t)$ is the vector displacement of any particle in solid; c_l and c_t are the velocities of a longitudinal and a transverse ultrasonic wave, respectively.

We shall begin by discussing a plane longitudinal elastic wave with condition $\mathrm{curl}\,\vec{u}(\vec{r},t) = 0$ and a plane transverse elastic wave with condition $\mathrm{div}\,\vec{u}(\vec{r},t) = 0$ in an infinite isotropic medium. In this respect, in direction of vector $\vec{r}$ can be propagated two transverse and one longitudinal elastic waves. The vector displacement $\vec{u}(\vec{r},t)$ is the sum of the vector displacements of a longitudinal $u_l(\vec{r},t)$ and of a transverse ultrasonic wave $u_t(\vec{r},t)$:

$$\vec{u}(\vec{r},t) = \vec{u}_l(\vec{r},t) + \vec{u}_t(\vec{r},t) \qquad (2)$$

where $\vec{u}_l(\vec{r},t)$ and $\vec{u}_t(\vec{r},t)$ are perpendicular with each other or $\vec{u}_l(\vec{r},t)\cdot\vec{u}_t(\vec{r},t) = 0$.

In turn, the equations of motion for a longitudinal and a transverse elastic wave take the form of the wave-equations:

$$\nabla^2\vec{u}_l(\vec{r},t) - \frac{1}{c_l^2}\frac{\partial^2\vec{u}_l(\vec{r},t)}{\partial t^2} = 0 \qquad (3)$$

$$\nabla^2\vec{u}_t(\vec{r},t) - \frac{1}{c_t^2}\frac{\partial^2\vec{u}_l(\vec{r},t)}{\partial t^2} = 0. \qquad (4)$$

It is well known, in quantum mechanics, a matter wave is determined by electromagnetic wave-particle duality or de Broglie wave of matter [4]. We argue that in an analogous manner, we may apply the elastic wave-particle duality. This reasoning allows us to present a model of elastic field as the Bose-gas consisting of the Sound Bose-particles with spin 1 and non-zero rest masses, which are interacting with each other. In this respect, we may express the vector displacement of a longitudinal ultrasonic wave $u_l(\vec{r},t)$ via the second quantization vector wave functions of one Sound Boson of the longitudinal wave. In analogy manner, vector displacement of a transverse ultrasonic waves $u_t(\vec{r},t)$ is expressed via the second quantization vector wave functions of one Sound Boson of the transverse wave:

$$\vec{u}_l(\vec{r},t) = u_l\Big(\phi(\vec{r},t) + \phi^+(\vec{r},t)\Big) \tag{5}$$

and

$$\vec{u}_t(\vec{r},t) = u_t\Big(\psi(\vec{r},t) + \psi^+(\vec{r},t)\Big) \tag{6}$$

where u_l and u_t are, respectively, the norm coefficients for longitudinal and transverse waves; $\vec{\phi}(\vec{r},t)$ and $\vec{\phi}^+(\vec{r},t)$ are, respectively, the second quantization wave vector functions for "creation" and "annihilation" of one Sound-Particle of the longitudinal wave, in point of coordinate $\vec{r}$ and time t whose direction $\vec{l}$ is directed toward to wave vector $\vec{k}$; $\vec{\psi}(\vec{r},t)$ and $\vec{\psi}^+(\vec{r},t)$ are, respectively, the second quantization wave vector functions for "creation" and "annihilation" of one Sound-Particle of the transverse wave, in point of coordinate $\vec{r}$ and time t, whose direction $\vec{t}$ is perpendicular to the wave vector $\vec{k}$:

$$\vec{\phi}(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{k},\sigma}\vec{a}_{\vec{k},\sigma}e^{i(\vec{k}\vec{r}-kc_lt)} \tag{7}$$

$$\vec{\phi}^+(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{k},\sigma}\vec{a}^+_{\vec{k},\sigma}e^{-i(\vec{k}\vec{r}-kc_lt)} \tag{8}$$

and

$$\vec{\psi}(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma}e^{i(\vec{k}\vec{r}+-kc_lt)} \tag{9}$$

$$\vec{\psi}^+(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{k},\sigma}\vec{b}^+_{\vec{k},\sigma}e^{-i(\vec{k}\vec{r}-kc_lt)} \tag{10}$$

with condition

$$\begin{aligned}&\int\phi^+(\vec{r},\sigma)\phi(\vec{r},\sigma)dV + \int\psi^+(\vec{r},\sigma)\psi(\vec{r},\sigma)dV = \\ &= n_o + \sum_{\vec{k}\neq0,\sigma}\hat{a}^+_{\vec{k},\sigma}\hat{a}_{\vec{k},\sigma} + \sum_{\vec{k}\neq0,\sigma}\hat{b}^+_{\vec{k},\sigma}\hat{b}_{\vec{k},\sigma} = \hat{n}\end{aligned} \tag{11}$$

where $\vec{a}^+_{\vec{k},\sigma}$ and $\vec{a}_{\vec{k},\sigma}$ are, respectively, the Bose vector-operators of creation and annihilation for one free longitudinal Sound Particle with spin 1, described by a vector $\vec{k}$ whose direction coincides with the direction $\vec{l}$ of the longitudinal wave; $\vec{b}^+_{\vec{k},\sigma}$ and $\vec{b}_{\vec{k},\sigma}$ are, respectively, the Bose vector-operators of creation and annihilation for one free transverse Sound Particles with spin 1, described by a vector $\vec{k}$ which is perpendicular to the direction $\vec{t}$ of the transverse wave; $\hat{n}$ is the operator of total number of the Sound Particles; $\hat{n}_0 = n_{0,l} + n_{0,t}$ is the total number of Sound Particles in the condensate level with wave vector $\vec{k} = 0$ which consists of a number of Sound Particles $n_{0,l}$ of the longitudinal wave and a number of Sound Particles $n_{0,t}$ of the transverse wave.

The energies of longitudinal $\frac{\hbar^2k^2}{2m_l}$ and transverse $\frac{\hbar^2k^2}{2m_t}$ free Sound Particles have the masses of Sound Particles m_l and m_t and the value of its spin z-component $\sigma = 0;\pm1$. In this respect, the vector-operators $\vec{a}^+_{\vec{k},\sigma}$, $\vec{a}_{\vec{k},\sigma}$ and $\vec{b}^+_{\vec{k},\sigma}$, $\vec{b}_{\vec{k},\sigma}$ satisfy the Bose commutation relations as:

$$\Big[\hat{a}_{\vec{k},\sigma},\hat{a}^+_{\vec{k}',\sigma'}\Big] = \delta_{\vec{k},\vec{k}'}\cdot\delta_{\sigma,\sigma'}$$

$$[\hat{a}_{\vec{k},\sigma},\hat{a}_{\vec{k}',\sigma'}] = 0$$

$$[\hat{a}^+_{\vec{k},\sigma},\hat{a}^+_{\vec{k}',\sigma'}] = 0$$

and

$$\Big[\hat{b}_{\vec{k},\sigma},\hat{b}^+_{\vec{k}',\sigma'}\Big] = \delta_{\vec{k},\vec{k}'}\cdot\delta_{\sigma,\sigma'}$$

$$[\hat{b}_{\vec{k},\sigma},\hat{b}_{\vec{k}',\sigma'}] = 0$$

$$[\hat{b}^+_{\vec{k},\sigma},\hat{b}^+_{\vec{k}',\sigma'}] = 0$$

Thus, as we see, the vector displacements of a longitudinal $\vec{u}_l$ and of a transverse $\vec{u}_t$ ultrasonic wave satisfy the wave-equations of (3) and (4) and have the forms:

$$\vec{u}_l(\vec{r},t) = u_{\vec{0},l} + \frac{u_l}{\sqrt{V}}\sum_{\vec{k}\neq0,\sigma}\left(\vec{a}_{\vec{k},\sigma}e^{i(\vec{k}\vec{r}-kc_lt)} + \vec{a}^+_{\vec{k},\sigma}e^{-i(\vec{k}\vec{r}-kc_lt)}\right) \tag{12}$$

and

$$\vec{u}_t(\vec{r},t) = u_{\vec{0},t} + \frac{u_t}{\sqrt{V}}\sum_{\vec{k}\neq0,\sigma}\left(\vec{b}_{\vec{k},\sigma}e^{i(\vec{k}\vec{r}-kc_lt)} + \vec{b}^+_{\vec{k},\sigma}e^{-i(\vec{k}\vec{r}-kc_lt)}\right). \tag{13}$$

While investigating superfluid liquid, Bogoliubov [5] separated the atoms of helium in the condensate from those atoms, filling states above the condensate. In an analogous manner, we may consider the vector operators $\hat{a}_0 = \vec{l}\sqrt{n_{0,l}}$, $\hat{b}_0 = \vec{t}\sqrt{n_{0,t}}$ and $\hat{a}^+_0 = \vec{l}\sqrt{n_{0,l}}$, $\hat{b}^+_0 = \vec{t}\sqrt{n_{0,t}}$ as c-numbers (where $\vec{l}$ and $\vec{t}$ are the unit vectors in the direction of the longitudinal and transverse elastic fields, respectively, and also $\vec{l}\cdot\vec{t} = 0$) within the approximation of a macroscopic number of Sound Particles in

the condensate $n_{0,l} \gg 1$ and $n_{0,t} \gg 1$. This assumptions lead to a broken Bose-symmetry law for Sound Particles of longitudinal and transverse waves in the condensate. In fact, we may state that if a number of Sound Particles of longitudinal and transverse waves fills a condensate level with the wave vector $\vec{k} = 0$, then they reproduce the constant displacements $\vec{u}_{0,l} = \frac{2u_l\vec{e}\sqrt{n_{0,l}}}{\sqrt{V}}$ and $\vec{u}_{0,t} = \frac{2u_t\vec{e}\sqrt{n_{0,t}}}{\sqrt{V}}$.

In this context, we may emphasize that the Bose vector operators $\vec{a}^{+}_{\vec{k},\sigma}$, $\vec{a}_{\vec{k},\sigma}$ and $\vec{b}^{+}_{\vec{k},\sigma}$ and $\vec{b}_{\vec{k},\sigma}$ communicate with each other because the vector displacements of a longitudinal $\vec{u}_l(\vec{r},t)$ and a transverse ultrasonic wave $\vec{u}_t(\vec{r},t)$ are independent, and in turn, satisfy to the Bose commutation relation $[\vec{u}_l(\vec{r},t), \vec{u}_t(\vec{r},t)] = 0$.

Now, we note that quantization of elastic field means that this field operator does not commute with its momentum density. Taking the commutators gives

$$\left[\vec{u}_l(\vec{r},t), \vec{p}_l(\vec{r}',t)\right] = i\hbar\delta^3_{\vec{r}-\vec{r}'} \tag{14}$$

and

$$\left[\vec{u}_t(\vec{r},t), \vec{p}_t(\vec{r}',t)\right] = i\hbar\delta^3_{\vec{r}-\vec{r}'} \tag{15}$$

where the momentums of the longitudinal and transverse waves are defined as

$$\vec{p}_l(\vec{r},t) = \rho_l(\vec{r})\frac{\partial\vec{u}_l(\vec{r},t)}{\partial t} \tag{16}$$

and

$$\vec{p}_t(\vec{r},t) = \rho_t(\vec{r})\frac{\partial\vec{u}_t(\vec{r},t)}{\partial t} \tag{17}$$

where $\rho_l(\vec{r})$ and $\rho_t(\vec{r})$ are, respectively, the mass densities of longitudinal and transverse Sound Particles in the coordinate space, which are presented by the equations

$$\rho_l(\vec{r}) = \rho_{0,l} + \sum_{\vec{k}\neq 0}\rho_l(\vec{k})e^{i\vec{k}\vec{r}} \tag{18}$$

and

$$\rho_t(\vec{r}) = \rho_{0,t} + \sum_{\vec{k}\neq 0}\rho_t(\vec{k})e^{i\vec{k}\vec{r}}. \tag{19}$$

The total mass density $\rho(\vec{r})$ is

$$\rho(\vec{r}) = \rho_0 + \sum_{\vec{k}\neq 0}\rho_l(\vec{k})e^{i\vec{k}\vec{r}} + \sum_{\vec{k}\neq 0}\rho_t(\vec{k})e^{i\vec{k}\vec{r}} \tag{20}$$

where $\rho_l(\vec{k})$ and $\rho_t(\vec{k})$ are, respectively, the fluctuations of the mass densities of the longitudinal and transverse Sound Particles which represent as the symmetrical function from wave vector $\vec{k}$ or $\rho_l(\vec{k}) = \rho_l(-\vec{k})$; $\rho_t(\vec{k}) = \rho_t(-\vec{k})$; $\rho_0 = \rho_{0,l} + \rho_{0,t}$ is the equilibrium density of Sound Particles.

Applying (12) and (13) to (16) and (17), and taking (18) and (19), we get

$$\vec{p}_l(\vec{r},t) = -\frac{ic_lu_l}{\sqrt{V}}\sum_{\vec{k}'}\sum_{\vec{k},\sigma}k\rho_l(\vec{k}')\Big(\vec{a}_{\vec{k},\sigma}e^{-ikc_lt} - \vec{a}^{+}_{-\vec{k},\sigma}e^{ikc_lt}\Big)e^{i(\vec{k}+\vec{k}')\vec{r}} \tag{21}$$

$$\vec{p}_t(\vec{r},t) = -\frac{ic_lu_l}{\sqrt{V}}\sum_{\vec{k}'}\sum_{\vec{k},\sigma}\rho_t(\vec{k}')k\Big(\vec{b}_{\vec{k},\sigma}e^{-ikc_lt} - \vec{b}^{+}_{-\vec{k},\sigma}e^{ikc_lt}\Big)e^{i(\vec{k}+\vec{k}')\vec{r}} \tag{22}$$

Application of (12), (21) and (13), (22) to (14) and (15), and taking the Bose commutation relations presented above, we obtain

$$\left[\vec{u}_l(\vec{r},t), \vec{p}_l(\vec{r}',t)\right] = \frac{2iu_l^2c_l}{V}\sum_{\vec{k}}k\rho_l(\vec{k})e^{i\vec{k}(\vec{r}-\vec{r}')} \tag{23}$$

and

$$\left[\vec{u}_t(\vec{r},t), \vec{p}_t(\vec{r}',t)\right] = \frac{2iu_t^2c_t}{V}\sum_{\vec{k}}k\rho_t(\vec{k})e^{i\vec{k}(\vec{r}-\vec{r}')} \tag{24}$$

The right sides of Eqs. (14) and (23) as well as Eqs. (15) and (24) coincide when

$$\rho_l(\vec{k}) = \frac{\hbar}{2ku_l^2c_l} \tag{25}$$

and

$$\rho_t(\vec{k}) = \frac{\hbar}{2ku_t^2c_t} \tag{26}$$

by using

$$\frac{1}{V}\sum_{\vec{k}}e^{i\vec{k}(\vec{r}-\vec{r}')} = \delta^3_{\vec{r}-\vec{r}'}$$

3 Sound-Particles and Phonons

The Hamiltonian operator $\hat{H}$ of the system, consisting of the vibrated Fermi-atoms with mass M, is represented by the following form

$$\hat{H} = \hat{H}_l + \hat{H}_t \tag{27}$$

where

$$\hat{H}_l = \frac{MN}{2V}\int\left(\frac{\partial\vec{u}_l}{\partial t}\right)^2dV + \frac{NM\Omega_l^2}{2V}\int(\vec{u}_l)^2dV \tag{28}$$

and

$$\hat{H}_t = \frac{MN}{2V}\int\left(\frac{\partial\vec{u}_t}{\partial t}\right)^2dV + \frac{NM\Omega_t^2}{2V}\int(\vec{u}_t)^2dV \tag{29}$$

with Ω_l and Ω_t which are, respectively, the natural frequencies of the atom by action of longitudinal and transverse elastic waves.

To find the Hamiltonian operator $\hat{H}$ of the system, we use the framework of Dirac [6] for the quantization of electromagnetic field:

$$\frac{\partial \vec{u}_l(\vec{r},t)}{\partial t} = -\frac{ic_l u_l}{\sqrt{V}} \sum_{\vec{k},\sigma} k\left(\vec{a}_{\vec{k},\sigma} e^{-ikc_l t} - \vec{a}^{+}_{-\vec{k},\sigma} e^{ikc_l t}\right) e^{i\vec{k}\vec{r}} \quad (30)$$

and

$$\frac{\partial \vec{u}_t(\vec{r},t)}{\partial t} = -\frac{ic_t u_t}{\sqrt{V}} \sum_{\vec{k},\sigma} k\left(\vec{b}_{\vec{k},\sigma} e^{-ikc_t t} - \vec{b}^{+}_{-\vec{k},\sigma} e^{ikc_t t}\right) e^{i\vec{k}\vec{r}} \quad (31)$$

which by substituting into (28) and (29) using (12) and (13), we obtain the reduced form for the Hamiltonian operators $\hat{H}_l$ and $\hat{H}_t$:

$$\hat{H}_l = \sum_{\vec{k},\sigma}\left[\left(\frac{MNu_l^2c_l^2k^2}{V} + \frac{MNu_l^2\Omega_l^2}{V}\right)\vec{a}^{+}_{\vec{k},\sigma}a_{\vec{k},\sigma} - \left(\frac{MNu_l^2c_l^2k^2}{V} - \frac{MNu_l^2\Omega_l^2}{V}\right)\left(\vec{a}_{-\vec{k},\sigma}\vec{a}_{\vec{k},\sigma} + \vec{a}^{+}_{\vec{k},\sigma}\vec{a}^{+}_{-\vec{k},\sigma}\right)\right] \quad (32)$$

and

$$\hat{H}_t = \sum_{\vec{k},\sigma}\left[\left(\frac{MNu_t^2c_t^2k^2}{V} + \frac{MNu_t^2\Omega_t^2}{V}\right)\vec{a}^{+}_{\vec{k},\sigma}a_{\vec{k},\sigma} - \left(\frac{MNu_t^2c_t^2k^2}{V} - \frac{MNu_l^2\Omega_t^2}{V}\right)\left(\vec{a}_{-\vec{k},\sigma}\vec{a}_{\vec{k},\sigma} + \vec{a}^{+}_{\vec{k},\sigma}\vec{a}^{+}_{-\vec{k},\sigma}\right)\right] \quad (33)$$

where u_l and u_t are defined by the first term in right side of (32) and (33) which represent as the kinetic energies of longitudinal Sound Particle $\frac{\hbar^2k^2}{2m_l}$ and transverse Sound Particles $\frac{\hbar^2k^2}{2m_t}$. Therefore, u_l and u_l are found, if we suggest:

$$\frac{MNu_l^2c_l^2k^2}{V} = \frac{\hbar^2k^2}{2m_l} \quad (34)$$

and

$$\frac{MNu_t^2c_t^2k^2}{V} = \frac{\hbar^2k^2}{2m_t} \quad (35)$$

which in turn determine

$$u_l = \frac{\hbar}{c_l\sqrt{2m_l\rho}}$$

and

$$u_t = \frac{\hbar}{c_t\sqrt{2m_t\rho}}$$

where $\rho = \frac{MN}{V}$ is the density of solid.

$$\hat{H}_l = \sum_{\vec{k},\sigma}\left[\left(\frac{\hbar^2k^2}{2m_l} + \frac{\hbar^2\Omega_l^2}{2m_lc_l^2}\right)\vec{a}^{+}_{\vec{k},\sigma}a_{\vec{k},\sigma} + \frac{U_{\vec{k},l}}{2}\left(\vec{a}_{-\vec{k},\sigma}\vec{a}_{\vec{k},\sigma} + \vec{a}^{+}_{\vec{k},\sigma}\vec{a}^{+}_{-\vec{k},\sigma}\right)\right] \quad (36)$$

and

$$\hat{H}_t = \sum_{\vec{k}}\left[\left(\frac{\hbar^2k^2}{2m} + \frac{\hbar^2\Omega_t^2}{2m_tc_t^2}\right)\vec{b}^{+}_{\vec{k},\sigma}b_{\vec{k},\sigma} + \frac{U_{\vec{k},t}}{2}\left(\vec{b}_{-\vec{k},\sigma}\vec{b}_{\vec{k},\sigma} + \vec{b}^{+}_{\vec{k},\sigma}\vec{b}^{+}_{-\vec{k},\sigma}\right)\right] \quad (37)$$

$U_{\vec{k},l}$ and $U_{\vec{k},t}$ are the interaction potentials between identical Sound Particles.

In analogous manner, as it was done in letter [7] regarding the quantization of the electromagnetic field, the boundary wave numbers $k_l = \frac{\Omega_l}{c_l}$ for the longitudinal elastic field and $k_t = \frac{\Omega_t}{c_t}$ for the transverse one are determined by suggestion that identical Sound Particles interact with each other by the repulsive potentials $U_{\vec{k},l}$ and $U_{\vec{k},t}$ in wave vector space

$$U_{\vec{k},l} = -\frac{\hbar^2k^2}{2m_l} + \frac{\hbar^2\Omega_l^2}{2m_lc_l^2} > 0$$

and

$$U_{\vec{k},t} = -\frac{\hbar^2k^2}{2m_t} + \frac{\hbar^2\Omega_t^2}{2m_tc_t^2} > 0$$

As results, there are two conditions for wave numbers of longitudinal $k < k_l$ and transverse $k < k_t$ Sound Particles which are provided by property of the model of hard spheres [8]. Indeed, there is a request of presence of repulsive potential interaction between identical kind of particles (recall S-wave repulsive pseudopotential interaction between atoms in the superfluid liquid ^{4}He in the model of hard spheres [8]).

On the other hand, it is well known that at absolute zero $T = 0$, the Fermi atoms fill the Fermi sphere in momentum space. As it is known, the total numbers of the Fermi atoms with opposite spins are the same, therefore, the Fermi wave number k_f is determined by a condition:

$$\frac{V}{2\pi^2}\int_0^{k_f} k^2dk = \frac{N}{2} \quad (38)$$

where N is the total number of Fermi-atoms in the solid. This reasoning together with the model of hard spheres claims the important condition as introduction the boundary wave number $k_f = \left(\frac{3\pi^2N}{V}\right)^{\frac{1}{3}}$ coinciding with k_l and k_t. Thus we claim that all Fermi atoms had one natural wavelength

$$\lambda_0 = \frac{2\pi}{k_f} = \frac{2\pi}{k_l} = \frac{2\pi}{k_t} \quad (39)$$

This approach is a similar to the Einstein model of solid where he suggested that all atoms have the same natural frequencies.

Now, to evaluate of the energy levels of the operator $\hat{H}_l$ (36) and $\hat{H}_t$ (37) in diagonal form, we use a new transformation of the vector-Bose-operators presented in [6]:

$$\vec{a}_{\vec{k},\sigma} = \frac{\vec{c}_{\vec{k},\sigma} + L_{\vec{k}}\vec{c}^{+}_{-\vec{k},\sigma}}{\sqrt{1 - L^2_{\vec{k}}}} \quad (40)$$

and

$$\vec{b}_{\vec{k},\sigma} = \frac{\vec{d}_{\vec{k},\sigma} + M_{\vec{k}}\vec{d}^{+}_{-\vec{k},\sigma}}{\sqrt{1 - M^2_{\vec{k}}}} \tag{41}$$

where $L_{\vec{k}}$ and $M_{\vec{k}}$ are, respectively, the real symmetrical functions of a wave vector $\vec{k}$. Consequently:

$$\hat{H}_l = \sum_{k<k_f,\sigma} \varepsilon_{\vec{k},l}\vec{c}^{+}_{\vec{k},\sigma}\vec{c}_{\vec{k},\sigma} \tag{42}$$

and

$$\hat{H}_t = \sum_{k<k_f,\sigma} \varepsilon_{\vec{k},t}\vec{d}^{+}_{\vec{k},\sigma}\vec{d}_{\vec{k},\sigma}. \tag{43}$$

Hence, we infer that the Bose-operators $\vec{c}^{+}_{\vec{k},\sigma}$, $\vec{c}_{\vec{k},\sigma}$ and $\vec{d}^{+}_{\vec{k},\sigma}$, $\vec{d}_{\vec{k},\sigma}$ are, respectively, the vector of "creation" and the vector of "annihilation" operators of longitudinal and transverse phonons with spin 1 and having the energies:

$$\varepsilon_{\vec{k},l} = \sqrt{\left(\frac{\hbar^2k^2}{2m_l} + \frac{\hbar^2\Omega_l^2}{2m_lc_l^2}\right)^2 - \left(\frac{\hbar^2k^2}{2m_l} - \frac{\hbar^2\Omega_l^2}{2m_lc_l^2}\right)^2} = \hbar kc_l \tag{44}$$

and

$$\varepsilon_{\vec{k},t} = \sqrt{\left(\frac{\hbar^2k^2}{2m_t} + \frac{\hbar^2\Omega_t^2}{2m_tc_t^2}\right)^2 - \left(\frac{\hbar^2k^2}{2m_t} - \frac{\hbar^2\Omega_l^2}{2m_tc_t^2}\right)^2} = \hbar kc_t \tag{45}$$

where the mass of longitudinal Sound Particle equals to

$$m_l = \frac{\hbar\Omega_l}{c_l^2} \tag{46}$$

but the mass of transverse Sound Particle is

$$m_t = \frac{\hbar\Omega_t}{c_t^2}. \tag{47}$$

Thus, we may state that there are two different Sound Particles with masses m_l and m_t which correspond to the longitudinal and transverse waves.

4 Thermodynamic property of solid

Now, we demonstrate that the herein presented theory leads to same results which were obtained by Debye in his theory investigating the thermodynamic properties of solids. So that, at the statistical equilibrium, the average energy of solid equals to

$$\overline{H} = \sum_{k<k_f,\sigma} \varepsilon_{\vec{k},l}\overline{\vec{c}^{+}_{\vec{k},\sigma}\vec{c}_{\vec{k},\sigma}} + \sum_{k<k_f,\sigma} \varepsilon_{\vec{k},t}\overline{\vec{d}^{+}_{\vec{k},\sigma}\vec{d}_{\vec{k},\sigma}} \tag{48}$$

where $\overline{\hat{c}^{+}_{\vec{k},\sigma}\hat{c}_{\vec{k},\sigma}}$ and $\overline{\hat{d}^{+}_{\vec{k},\sigma}\hat{d}_{\vec{k},\sigma}}$ are, respectively, the average number of phonons with the wave vector $\vec{k}$ corresponding to the longitudinal and transverse fields at temperature T:

$$\overline{\hat{c}^{+}_{\vec{k},\sigma}\hat{c}_{\vec{k},\sigma}} = \frac{1}{e^{\frac{\varepsilon_{\vec{k},l}}{kT}} - 1}$$

and

$$\overline{\hat{d}^{+}_{\vec{k},\sigma}\hat{d}_{\vec{k},\sigma}} = \frac{1}{e^{\frac{\varepsilon_{\vec{k},t}}{kT}} - 1}.$$

Thus, at thermodynamic limit, the average energy of solid may rewritten down as

$$\overline{H} = \frac{3Vk^4T^4}{2\pi^2\hbar^3c_l^3}\int_0^{\frac{\Theta_l}{T}} \frac{x^3dx}{e^x - 1} + \frac{3Vk^4T^4}{2\pi^2\hbar^3c_t^3}\int_0^{\frac{\Theta_t}{T}} \frac{x^3dx}{e^x - 1} \tag{49}$$

where $\Theta_l = \frac{\hbar k_f c_l}{k}$ and $\Theta_t = \frac{\hbar k_f c_t}{k}$ are, respectively, the characteristic temperatures for solid corresponding to longitudinal and transverse waves; k is the Boltzmann constant. In our theory we denote

$$\frac{1}{c_l^3} + \frac{1}{c_t^3} = \frac{2}{c^3}$$

where c is the average velocity of phonons with spin 1 in the given theory; $\Theta_B = \frac{\hbar k_f c}{k}$ is the new characteristic temperature.

Hence, we may note that the coefficient with number 3 must be appear before both integrals on the right side of equation (49) because it reflects that phonons of longitudinal and transverse waves have number 3 quantities of the value of spin z-component $\sigma = 0; \pm 1$. At $T \ll \Theta_l$ and $T \ll \Theta_t$, the equation (49) takes the form:

$$\overline{H} = \frac{3\pi^4}{5}\frac{NkT^4}{2}\left(\frac{1}{\Theta_l^3} + \frac{1}{\Theta_t^3}\right) \tag{50}$$

where $\int_0^\infty \frac{x^3dk}{e^x-1} = \frac{\pi^4}{15}$.

Thus, Eq.(50) may be rewritten as

$$\overline{H} \approx \frac{3\pi^4}{5}\frac{RT^4}{\Theta_B^3} \tag{51}$$

where $R = Nk$ is the gas constant. Hence, we may note that at $T \gg \Theta_l$ and $T \gg \Theta_t$, the equation (49) takes the form:

$$\overline{H} = 3RT. \tag{52}$$

In this context, the heat capacity is determined as

$$C_V = \left(\frac{d\overline{H}}{dT}\right)_V \tag{53}$$

which obviously, at $T \ll \Theta_l$ and $T \ll \Theta_t$, the equation (53) with (51) reflects the Debye law T^3 at low temperatures:

$$C_V \approx \frac{12\pi^4}{5}\frac{RT^3}{\Theta_B^3}. \tag{54}$$

But at high temperatures $T \gg \Theta_l$ and $T \gg \Theta_t$, the equation (53) with (52) recovers the Dulong-Petit law:

$$C_V \approx 3R. \tag{55}$$

Obviously, the average velocity of phonon c and new characteristic temperature Θ_B are differ from their definition in Debye theory because the average energy of solid in Debye theory is presented as

$$\overline{H}_D = \frac{3Vk^4T^4}{2\pi^2\hbar^3c_l^3}\int_0^{\frac{\Theta_l}{T}}\frac{x^3dx}{e^x-1} + \frac{3Vk^4T^4}{\pi^2\hbar^3c_t^3}\int_0^{\frac{\Theta_t}{T}}\frac{x^3dx}{e^x-1} \qquad (56)$$

where $\Theta_l = \frac{\hbar k_D c_l}{k}$ and $\Theta_t = \frac{\hbar k_D c_t}{k}$ are, respectively, the characteristic temperatures for solid corresponding to one longitudinal and two transverse waves:

$$\frac{1}{c_l^3} + \frac{2}{c_t^3} = \frac{3}{v_0^3} \qquad (57)$$

where v_0 is the average velocity of spinless phonons in Debye theory; $k_D = \left(\frac{6\pi^2N}{V}\right)^{\frac{1}{3}}$ is the Debye wave number; $\Theta_D = \frac{\hbar k_D v_0}{k}$ is the Debye characteristic temperature which is

$$\frac{1}{\Theta_l^3} + \frac{2}{\Theta_t^3} = \frac{3}{\Theta_D^3} \qquad (58)$$

As we see the average energy of solid $\overline{H}_D$ in (56) is differ from one in (49) by coefficient 2 in ahead of second term in right side of Eq.(56) (which is connected with assumption of presence two transverse waves), as well as introduction of Debye wave number k_D. So that due to definition of the average velocity v_0 of spinless phonons by (57), Debye may accept a phonon as spinless quasiiparticle.

5 Concussion

Thus, in this letter, we propose new model for solids which is different from the well-known models of Einstein and Debye because: 1), we suggest that the atoms are the Fermi particles which are absent in the Einstein and Debye models; 2), we consider the stimulated oscillation of atoms by action of longitudinal and transverse waves in the solid. The elastic waves stimulate the vibration of the fermion-atoms with one natural wavelength, we suggested that atoms have two independent natural frequencies corresponding to a longitudinal and a transverse wave, due to application of the principle of the elastic wave-particle duality, the model of hard spheres and considering the atoms as the Fermi particles. In accordance to this reasoning, there is an appearance of a cut off in the energy spectrum of phonons; 3), In our model, we argue that the photons have spin 1 which is different from models presented by Einstein and Debye. On the other hand, we suggest that only one longitudinal and one transverse wave may be excited in the lattice of the solid which is different from Debye who suggested a presence of two sorts of transverse waves.

The quantization of the elastic wave by our theory leads to a view of the lattice as the diffraction picture. Within our theory, the mass density $\rho(\vec{r})$ in coordinate space, due to substituting $\rho_l(\vec{r},t)$ and $\rho_l(\vec{r},t)$ from (25) and (26) into (20), represents as

$$\rho(\vec{r}) = \rho_0 + \frac{8\pi\hbar k_f^2}{u_l^2 c_l}\left(\frac{\sin k_f r}{k_f r}\right)^2 + \frac{8\pi\hbar k_f^2}{u_t^2 c_t}\left(\frac{\sin k_f r}{k_f r}\right)^2 \qquad (59)$$

which implies that the lattice has the diffraction picture.

Now, we try to estimate the masses of the Sound Particles in substance as Aluminium Al. In this respect, we use of (46) and (47) with introducing of the Fermi momentum $p_f = \hbar k_f = \frac{\hbar\Omega_l}{c_l} = \frac{\hbar\Omega_t}{c_t}$, for instance, for such material as Al with $c_l = 6.26\cdot 10^3\frac{m}{sec}$ and $c_t = 3.08\cdot 10^3\frac{m}{sec}$ at room temperature [9], and $p_f = 1.27\cdot 10^{-24}\frac{kg\cdot m}{sec}$ we may estimate $m_l = \frac{p_f}{c_l} = 2\cdot 10^{-28}kg$ and $m_t = \frac{p_f}{c_t} = 4\cdot 10^{-28}kg$.

It is well known that the mass of atom Al is $M = 10^{-25}kg$ which is around 500 time more in regard to the masses of Sound Particles.

In this context, we remark that the new characteristic temperature Θ_B almost coincide with the Debye temperature Θ_D. Indeed, by our theory for Al:

$$\Theta_B = \frac{2^{\frac{1}{3}}p_f c_l}{k\left(1+\frac{c_l^3}{c_t^3}\right)^{\frac{1}{3}}} \approx 400K$$

but Debye temperature equals to $\Theta_D = 418K$.

Acknowledgements

We are particularly grateful to Dr. Andreas Ries for valuable scientific support and for help with the English.

Submitted on October 20, 2010 / Accepted on November 9, 2010

References

1. Einstein A. Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Waerme. *Annalen der Physik*, 1907, v. 22, 180–190.
2. Debye P. Zur Theorie der spezifischen Waerme. *Annalen der Physik*, 1912, v. 39, no. 4, 789–839.
3. Landau L. D., Lifshiz E. M. Theory of Elasticity. *Theoretical Physics* 1987, v. V11, 124–127.
4. de Broglie L. Researches on the quantum theory. *Annalen der Physik*, 1925, v. 3, 22–32.
5. Bogoliubov N. N. On the theory of superfludity. *Journal of Physics (USSR)*, 1947, v. 11, 23–32.
6. Dirac P. A. M. The Principles of Quantum Mechanics. Clarendon Press, Oxford, 1958.
7. Minasyan V. N., Samoilov V. N. Two Type Surface Polaritons Excited into Nanoholes in Metal Films. *Progress in Physics*, 2010, v. 2, 3–6.
8. Huang K. Statistical Mechanics. John Wiley, New York, 1963.
9. Kittel Ch., Introduction to Solid State Physics. John Wiley, New York, 1990.

Applying Adjacent Hyperbolas to Calculation of the Upper Limit of the Periodic Table of Elements, with Use of Rhodium

Albert Khazan

E-mail: albkhazan@gmail.com

In the earlier study (Khazan A. Upper Limit in Mendeleev's Periodic Table — Element No. 155. 2nd ed., Svenska fysikarkivet, Stockholm, 2010) the author showed how Rhodium can be applied to the hyperbolic law of the Periodic Table of Elements in order to calculate, with high precision, all other elements conceivable in the Table. Here we obtain the same result, with use of fraction linear functions (adjacent hyperbolas).

1 Introduction

In the theoretical deduction of the hyperbolic law of the Periodic Table of Elements [1], the main attention was focused onto the following subjects: the equilateral hyperbola with the central point at the coordinates (0; 0), its top, the real axis, and the line tangential to the normal of the hyperbola. All these were created for each element having the known or arbitrary characteristics. We chose the top of the hyperbolas, in order to describe a chemical process with use of Lagrange's theorem; reducing them to the equation $Y = K/X$ was made through the scaling coefficient 20.2895, as we have deduced.

The upper limit of the Table of Elements, which is the heaviest (last) element of the Table, is determined within the precision we determine the top of its hyperbola [1]. Therefore hyperbolas which are related to fraction linear functions were deduced. These hyperbolas are equilateral as well, but differ in the coordinates of their centre: $x = 0$, $y = 1$. To avoid possible mistakes in the future, the following terminology has been assumed: hyperbolas of the kind $y = k/x$ are referred to as *straight*; equilateral hyperbolas of the kind $y = (ax + b)/(cx + d)$ are referred to as *adjacent*. The latter ones bear the following properties: such a hyperbola intersects with the respective straight hyperbola at the ordinate $y = 0.5$ and the abscissa equal to the double mass of the element; the line $y = 0.5$ is the axis of symmetry for the arcs; the real and tangential lines of such hyperbolas meet each other; the normal of such a hyperbola is the real axis and the tangential line of another hyperbola of this kind.

The found common properties of the hyperbolas provided a possibility to use them for determination of the heaviest (last) element in another way than earlier.

2 Method of calculation

Once drawing straight hyperbolas for a wide range of the elements, according to their number from 1 to 99 in the Table of Elements, where the atomic masses occupy the scale from Hydrogen (1.00794) to Einsteinium (252), one can see that the real axis of each straight hyperbola is orthogonal to the real axis of the respective adjacent hyperbola, and they cross each other at the point $y = 0.5$.

Then we draw the intersecting lines from the origin of the adjacent hyperbolas (0; 1). The lines intersect the straight hyperbolas at two points, and also intersect the real axis and the abscissa axis where they intercept different lengths.

Connection to molecular mass of an element (expressed in the Atomic Units of Mass) differs between the abscissas of the lengths selected by the intersecting lines and the abscissas of transection of the straight and adjacent hyperbolas. Therefore, the line which is tangent to the straight in the sole point (102.9055; 205.811) is quite complicated. These coordinates mean the atomic mass of Rhodium and the half of the atomic mass of the heaviest (last) element of the Periodic Table.

The right side of the line can easily be described by the 4th grade polynomial equation. However the left side has a complicate form, where the maximum is observed at the light elements (Nitrogen, Oxygen) when lowered to (102.9055; 0) with the increase in atomic mass.

According to our calculation, the straight and adjacent hyperbolas were determined for Rhodium. The real axes go through the transecting points of the hyperbolas to the axis X and the line $Y = 1$, where they intercept the same lengths 411.622. This number differs for 0.009% from 411.66.

Thus, this calculation verified the atomic mass 411.66 of the heaviest element (upper limit) of the Periodic Table of Elements, which was determined in another way in our previous study [1].

3 Algorithm of calculation

The algorithm and results of the calculation without use of Rhodium were given in detail in Table 3.1 of the book [1]. The calculation is produced in six steps.

Step 1. The data, according to the Table of Elements, are written in columns 1, 2, 3.

Step 2. Square root is taken from the atomic mass of each element. Then the result transforms, through the scaling coefficient 20.2895, into the coordinates of the tops of straight hyperbolas along the real axis. To do it, the square root of the data of column 3 is multiplied by 20.2895 (column 4), then is divided by it (column 3).

Step 3. We draw transecting lines from the centre (0; 1) to

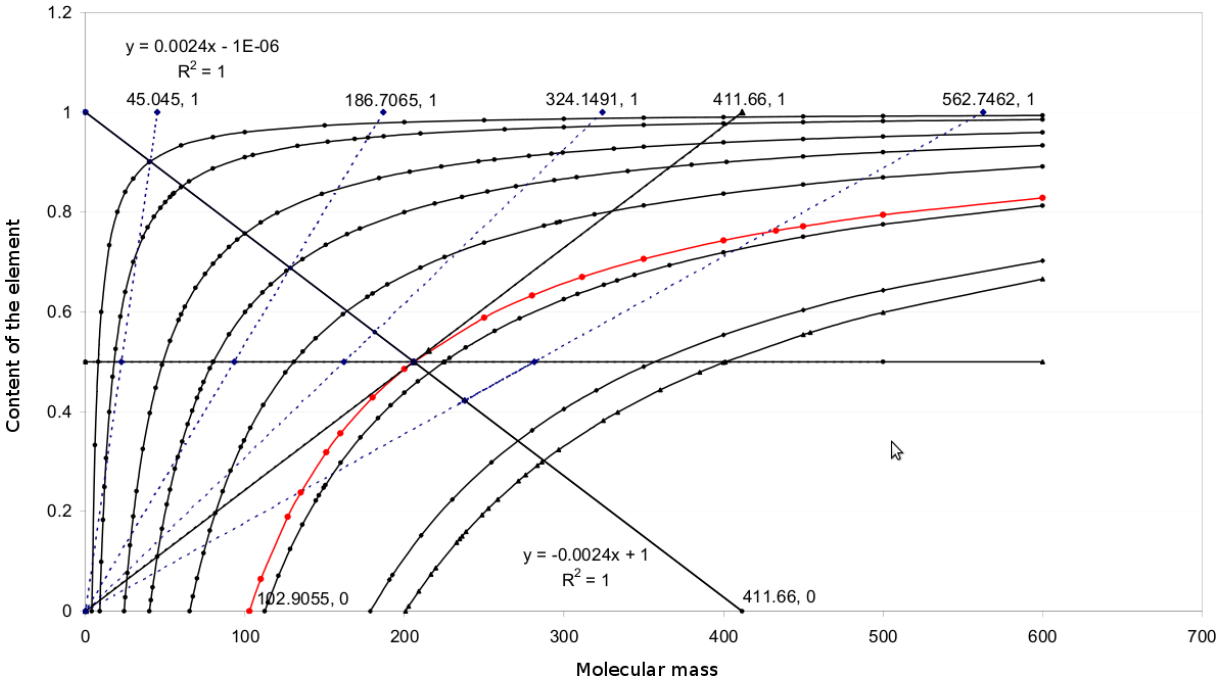


Fig. 1: Calculation with the centre at the point (0;0).

the transections with the line $y = 0.5$, with the real axis at the point $(X_0; Y_0)$, and so forth up to the axis X. To determine the abscissa of the intersection points, we calculate the equation of a straight line of each element. This line goes through two points: the centre (0; 1) and a point located in the line $y = 0.5$ or in the axis X: $(X - 0)/(X_0 - 0) = (Y - 1)/(Y_0 - 1)$. For instance, consider Magnesium. After its characteristics substituted, we obtain the equation $(X - 0)/(100.0274 - 0) = (Y - 1)/(0.242983 - 1)$, wherefrom the straight line equation is obtained: $Y = 1 - 0.007568\,X$. Thus, the abscissa of the transecting line, in the line $y = 0.5$, is 66.0669 (column 6).

Step 4. We write, in column 7, the abscissas of the points of transection of the straight and adjacent hyperbolas. The abscissas are equal to the double atomic mass of the element under study.

Step 5. We look for the region, where the segment created by a hyperbola and its transecting line is as small as a point (of the hyperbola and its transecting line). To find the coordinates, we subtract the data of column 7 from the respective data of column 6. Then we watch where the transecting line meets the real axis. The result is given by column 8. Here we see that the numerical value of the segments increases, then falls down to zero, then increases again but according to another law.

Step 6. Column 9 gives tangent of the inclination angle of the straights determined by the equations, constructed for two coordinate points of each element: $Y = -KX + 1$, where K is the tangent of the inclination angle.

4 Using adjanced hyperbolas in the calculation

Because straight and adjacent hyperbolas are equilateral, we use this fact for analogous calculations with another centre, located in the point (0; 0). The result has been shown in Fig. 1. In this case X_0 remains the same, while the ordinate is obtained as difference between 1 and Y_0. The straight line equation is obtained between two points with use of the data of column 9, where tangent should be taken with the opposite sign. As a result, we obtain an adjacent hyperbola of Rhodium. For example, consider Calcium. We obtain $X_0 = 128.4471$, $Y_0 = 0.31202$ (the ordinate for the straight hyperbola of Calcium), and $Y_0 = 1 - 0.31202 = 0.68798$ (for the adjacent hyperbola). The straight line equation between these two points is $Y = 0.005356\,X$. Thus, we obtain $x = 186.7065$ under $y = 1$, and $x = 93.3508$ under $y = 0.5$.

The new calculations presented here manifest that determining the heaviest (last) element of the Periodic Table of Elements is **correct for both ways of calculation**: the way with use of Lagrange's theorem and the scaling coefficient [1], and also the current method of the hyperbolas adjacent to that of Rhodium (method of adjacent hyperbolas). As one can see, the calculation results obtained via these two methods differ only in thousand doles of percent.

Submitted on October 22, 2010 / Accepted on December 12, 2010

References

1. Khazan A. Upper Limit in Mendeleev's Periodic Table — Element No. 155, 2nd edition, Svenska fysikarkivet, Stockholm, 2010.

How Black Holes Violate the Conservation of Energy

Douglas L. Weller
Email: physics@dougweller.com

Black holes produce more energy than they consume thereby violating the conservation of energy and acting as perpetual motion machines.

1 Introduction

According to Stephen Hawking and Leonard Mlodinow [1]: "Because there is a law such as gravity, the Universe can and will create itself from nothing." Such views of gravity are usually attributed as being rooted in Einstein's general-relativistic space-time.

However, the field equations Einstein [2] used to describe the general-relativistic space-time are founded on the conservation of momentum and energy. How can a space-time derived based on the conservation of momentum and energy provide an ex nihilo source of energy sufficient to create a universe?

The answer is found in Karl Schwarzschild's solution [3] to the field equations, usually called the Schwarzschild metric. The Schwarzschild metric describes a gravitational field outside a spherical non-rotating mass. When the mass is compacted within its Schwarzschild radius it is commonly referred to as a black hole.

Herein the terms of the Schwarzschild metric are rearranged to display limits in the Schwarzschild metric that necessarily result from the conservation of momentum and energy. Then is shown how black holes violate the limits, acting as perpetual motion machines that produce more energy than they consume.

2 Expressing the Schwarzschild metric using velocities

In this section, the Schwarzschild metric is rearranged so as to be expressed using velocities measured with reference coordinates. This rearrangement, which appears as equation (8) at the end of this section, will make very clear the limits imposed within the Schwarzschild metric by the conservation of momentum and energy.

Einstein [4] originally expressed the principles of special relativity using velocities measured with reference coordinates. However, Einstein [2, Equations 47] expressed the field equations in more abstract terms, using tensors. Einstein was careful to show that the field equations, nevertheless, correspond to the conservation of momentum and energy [2, Equations 47a] and thus have a nexus to physical reality.

The Schwarzschild metric, as a solution to the field equations, also corresponds to the conservation of momentum and energy. Arrangement of the Schwarzschild metric as in (8) allows for an intuitive comprehension of exactly how momentum and energy is conserved.

For a compact mass M with a Schwarzschild radius R, the Schwarzschild metric is often expressed using reference space coordinates (r, θ, ϕ), coordinate time t and local time τ (often referred to as proper time τ), as

$$c^2 d\tau^2 = c^2 \left(1 - \frac{R}{r}\right) dt^2 - \frac{dr^2}{(1 - R/r)} - r^2 d\theta^2 - (r sin\theta)^2 d\phi^2. \quad (1)$$

The Schwarzschild metric as shown in (1) can be rearranged to form (8), as shown below. To obtain (8) from (1), begin by multiplying both sides of (1) by $\left(\frac{1}{dt}\right)^2$ yielding

$$\begin{aligned} c^2 \left(\frac{d\tau}{dt}\right)^2 &= c^2 \left(1 - \frac{R}{r}\right)\left(\frac{dt}{dt}\right)^2 - \frac{1}{1 - R/r}\left(\frac{dr}{dt}\right)^2 \\ &- r^2 \left(\frac{d\theta}{dt}\right)^2 - (r sin\theta)^2 \left(\frac{d\phi}{dt}\right)^2, \end{aligned} \quad (2)$$

which allows motion in all dimensions to be measured with respect to the reference coordinates (r, θ, ϕ, t). The terms of (2) can be rearranged as

$$\begin{aligned} c^2 &= c^2 \left(\frac{d\tau}{dt}\right)^2 + c^2 \frac{R}{r} + \frac{1}{1 - R/r}\left(\frac{dr}{dt}\right)^2 \\ &+ r^2 \left(\frac{d\theta}{dt}\right)^2 + (r sin\theta)^2 \left(\frac{d\phi}{dt}\right)^2. \end{aligned} \quad (3)$$

The terms in (3) can be grouped by defining three different velocities. A velocity through the three dimensions of curved space can be defined as

$$v_S = \sqrt{\frac{1}{1 - R/r}\left(\frac{dr}{dt}\right)^2 + r^2 \left(\frac{d\theta}{dt}\right)^2 + (r sin\theta)^2 \left(\frac{d\phi}{dt}\right)^2}. \quad (4)$$

A velocity of local time through a time dimension can be defined as

$$v_\tau = c \frac{d\tau}{dt}. \quad (5)$$

A gravitational velocity can be defined as

$$v_G = c \sqrt{\frac{R}{r}}. \quad (6)$$

Using the definitions in (4), (5) and (6), (3) reduces to

$$c^2 = v_\tau^2 + v_G^2 + v_S^2. \quad (7)$$

Equation (7) can be expressed using orthogonal vectors $\vec{v}_\tau$, $\vec{v}_G$ and $\vec{v}_S$ where $v_\tau = |\vec{v}_\tau|$, $v_G = |\vec{v}_G|$ and $v_S = |\vec{v}_S|$, and where

$$c = \left|\vec{v}_\tau + \vec{v}_G + \vec{v}_S\right|. \quad (8)$$

Equation (8) is mathematically equivalent to (1) and expresses the Schwarzschild metric as a relationship of vector velocities. The conservation of momentum and energy, as expressed in the Schwarzschild metric, requires that the magnitude of the sum of the velocities is always equal to the constant c. Before exploring the full implication of this relationship, the next section confirms that (8) conforms with what is predicted by special relativity.

3 Equation (8) and special relativity

In the previous section, the Schwarzschild metric in (1) has been rearranged as (8) to provide a more concrete picture of the relationships necessary for conservation of momentum and energy.

Here is confirmed (8) is in accord with the case of special relativity for unaccelerated motion.

When there is no acceleration and therefore no gravity field, $R = 0$ and thus according to (6), $v_G = 0$ so that (8) reduces to

$$c = \left|\vec{v}_\tau + \vec{v}_S\right|. \tag{9}$$

When $R = 0$,

$$v_{S,R=0} = \sqrt{\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2 + (rsin\theta)^2\left(\frac{d\phi}{dt}\right)^2}, \tag{10}$$

which expressed in Cartesian coordinates is the familiar form of velocity used in special relativity,

$$v_{S,R=0} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}. \tag{11}$$

Equation (9) accurately reproduces the relationship of velocity and time known from special relativity. As velocity v_S in the space dimensions increases, there is a corresponding decline in the velocity v_τ in the orthogonal time dimension. When velocity in the time dimension reaches its minimum value (i.e., $v_\tau = 0$) this indicates a maximum value (i.e., $v_S = c$) in the space dimensions has been reached.

Equation (9) can be rearranged to confirm it portrays exactly the relationship between coordinate time and local time that is known to occur in the case of special relativity. Specifically, from the relationship of the orthogonal vectors $\vec{v}_\tau$, and $\vec{v}_S$ in (9), it must be true that

$$c^2 = v_\tau^2 + v_S^2. \tag{12}$$

and thus from (5)

$$c^2 = c^2\left(\frac{d\tau}{dt}\right)^2 + v_S^2, \tag{13}$$

and therefore

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{v_S^2}{c^2}}, \tag{14}$$

which is a form of the well known Laplace factor indicating the relationship between local time and coordinate time for special relativity.

4 Equation (8) and limits imposed by the conservation of momentum and energy

The arrangement of the Schwarzschild metric in (8) allows for a more concrete explanation of the limitations inherent in the Schwarzschild metric that necessarily result from the conservation of momentum and energy.

The vector sum of $\vec{v}_\tau$, $\vec{v}_G$ and $\vec{v}_S$ establishes a maximum value of c for each individual vector velocity.

When $\vec{v}_\tau = 0$ and $\vec{v}_S = 0$, $\vec{v}_G$ reaches its maximum value of c. Gravitational velocity $\vec{v}_G$ cannot exceed its maximum value of c without violating (8).

According to the definition of v_G in (6), when $v_G = c$, then $r = R$. When $r < R$, then $v_G > c$; therefore, according to (8), $r < R$ never occurs. As shown by Weller [5], matter from space can never actually reach $r = R$, but if it could, it would go no farther. At $r = R$ and $v_G = c$, all motion through space stops ($\vec{v}_S = 0$) and local time stops ($\vec{v}_\tau = 0$, so $d\tau/dt = 0$). Without motion in time or space, matter cannot pass through radial location $r = R$.

This section has shown that because of the conservation of momentum and energy — as expressed by the Schwarzschild metric arranged as in (8) — matter from space cannot cross the Schwarzschild radius R to get to a location where $r < R$.

The following sections consider conservation of energy equivalence in the Schwarzschild metric and the result when energy conservation is not followed.

5 Apportionment of energy equivalence

Einstein [6] pioneered apportioning energy differently based on reference frames, using such an apportionment in his initial calculations deriving the value for the energy equivalence of a mass (i.e., $E = mc^2$).

This notion of apportionment of energy equivalence is a helpful tool in understanding the implications of violating the conservation of energy and momentum in the Schwarzschild metric. When considering apportionment of energy equivalence in the Schwarzschild metric, it is helpful to keep in mind how Einstein makes a distinction between "matter" and a "matterless" gravitational field defined by the field equations or by the Schwarzschild metric. According to the Einstein [2, p. 143], everything but the gravitation field is denoted as "matter". Therefore, matter when added to the matterless field includes not only matter in the ordinary sense, but the electromagnetic field as well.

How the Schwarzschild metric apportions energy equivalence can be understood from

$$c^2 = v_\tau^2 + c^2\frac{R}{r} + v_S^2. \tag{15}$$

which is (7) modified so as to replace v_G with its equivalent given in (6). Equation (15) is mathematically equivalent to (1), just rearranged to aid in the explanation of the apportionment of energy equivalence.

Equation (15) can be put into perhaps more familiar terms by introducing a particle of mass m into the gravitation field. The energy equivalence mc^2 of the mass m is apportioned according to (15) as

$$mc^2 = mv_\tau^2 + mc^2\frac{R}{r} + mv_S^2. \tag{16}$$

In order to provide insight into the nature of the gravitational energy component c^2R/r in (15) —which appears as mc^2R/r in (16) — the next section discusses briefly how this term came to reside in the Schwarzschild metric.

6 Schwarzschild's description of gravity

One of the issues Schwarzschild [3, see §4] faced when deriving the Schwarzschild metric was how to describe the effects of gravity. He chose to do so using a positive integration constant that depends on the value of the mass at the origin. As a result the Newtonian gravitational constant G appears in the Schwarzschild metric. In (1) the gravitational constant G appears as part of the definition of the Schwarzschild radius R. In both Newtonian physics and the Schwarzschild metric, the Schwarzschild radius (R) — the location where Newtonian escape velocity (i.e., v_G) is equal to c — is defined as

$$R = \frac{2GM}{c^2}. \tag{17}$$

When the Schwarzschild metric is arranged as in (15), gravitational energy component c^2R/r increases toward infinity as radial location r decreases toward zero. This suggests the location of an unlimited energy source within the Schwarzschild metric; however, total gravitational energy is limited by the requirement that energy be conserved, as illustrated by the hypothetical described in the next section.

7 A hypothetical illustrating the conservation of energy equivalence in the Schwarzschild metric

The total energy-equivalence of a system comprised of a mass M can be defined as

$$E_M = Mc^2, \tag{18}$$

where the energy of magnetic fields is included in M, or neglected. If a mass m is added to the system, the additional energy E added to the system as a result of the presence of mass m is also well known to be

$$E = mc^2. \tag{19}$$

Thus if the system consisting of mass M and mass m were dissolved into radiation, the total resulting energy would be equal to

$$E_M + E = Mc^2 + mc^2. \tag{20}$$

In order for the conservation of energy to be maintained in the system as a whole, any gravitational energy E_G or any energy from motion E_K that is added to the system as a result of the presence of mass m must be included as part of the additional energy E described in (19). Therefore, the additional energy E present in the system as a result of adding mass m can be expressed as

$$E = mc^2 = E_K + E_G + E_\tau, \tag{21}$$

where E_τ is the portion of energy E that is not represented by gravitational energy component E_G or motion energy component E_K.

Equation (21) is the apportionment of energy equivalence shown in (16). To confirm this, in (21) set $E_G = mc^2R/r$, $E_K = mv_S^2$ and $E_\tau = mv_\tau^2$ to obtain (16).

The apportionment of energy equivalence in (16) and (21) indicates why crossing the Schwarzschild radius R violates the conservation of energy. When the particle reaches the Schwarzschild radius R — i.e., $r = R$ — the entire energy equivalence of mass m, is consumed by the gravitation component, i.e., $E_G = mc^2R/R = mc^2$. There is no energy left for mass m to travel in time (i.e., $E_\tau = 0$) or in space (i.e., $E_K = 0$). Therefore at locations $r = R$, all motion in time and space must stop, preventing mass m from ever crossing the critical radius.

If mass m were from space to cross the Schwarzschild radius R, the gravitational energy component $E_G = mc^2R/r$ would exceed the total energy equivalence $E = mc^2$ violating the conservation of energy.

If the particle were allowed to reach $r = 0$, gravitational energy component $E_G = mc^2R/r$ would approach infinity before becoming undefined.

8 How black holes act as perpetual motion machines

A perpetual motion machine is a hypothetical machine that violates the conservation of energy by producing more energy than it consumes.

According to the conservation of momentum and energy described by the Schwarzschild metric, see (8) and (15), a particle can never from space cross the Schwarzschild radius R of a compact mass M.

When a black hole is formed from a compacting mass M, the last particle on the surface of the mass that reaches and crosses the Schwarzschild radius R violates (8). Every particle thereafter that from space crosses R violates (8).

Further, from (16), each particle of mass m that reaches a radial location $r < R$, produces an amount of gravitational energy ($E_G = mc^2R/r$) that is greater than its total energy equivalence mc^2, as can only happen in a perpetual motion machine. When a particle is allowed to approach and reach $r = 0$, the ultimate perpetual motion machine is created which from the finite energy equivalence mc^2 of the particle produces an unlimited amount of gravitational energy as the particle approaches $r = 0$.

9 Concluding Remarks

Describing the effects of gravity using a gravitational constant and violating the conservation of momentum and energy described by the Schwarzschild metric can hypothetically result in black holes that act as perpetual motion machines able to produce an unlimited amount of energy. However, the existence of such perpetual motion machines is not in accordance with the conservation of momentum and energy as expressed in Einstein's general-relativistic space-time.

Special mathematical calculations, including use of specially selected coordinates, have been used to explain how a particle can cross the Schwarzschild radius allowing black holes to form. Critiquing these mathematical calculations is beyond the scope of this short paper. The author has directly addressed some of this subject matter in a companion paper [5].

Submitted on November 16, 2010 / Accepted on December 15, 2010

References

1. Hawking S., Mlodinow L. The grand design. Bantom Books, New York, 2010, p. 180.
2. Einstein A. The foundation of the general theory of relativity. The Principle of Relativity. Dover Publications, New York, 1923. pp. 111–164.
3. Schwarzschild K. On the gravitational field of a mass point according to Einstein's theory. Translated by S. Antoci, A. Loinger, 1999, arXiv:Physics/9905030.
4. Einstein A. On the electrodynamics of moving bodies. The Principle of Relativity. Dover Publications, New York, 1923, pp. 37–65.
5. Weller D. Five fallacies used to link black holes to Einstein's relativistic space-time. *Progress in Physics*, 2011, v. 1, 93–97.
6. Einstein A. Does the inertia of a body depend upon its energy-content? The Principle of Relativity. Dover Publications, New York, 1923, pp. 69–71.

Five Fallacies Used to Link Black Holes to Einstein's Relativistic Space-Time

Douglas L. Weller
E-mail: physics@dougweller.com

For a particle falling radially toward a compact mass, the Schwarzschild metric maps local time to coordinate time based on radial locations reached by the particle. The mapping shows the particle will not cross a critical radius regardless of the coordinate used to measure time. Herein are discussed five fallacies that have been used to make it appear the particle can cross the critical radius.

1 Introduction

Einstein [1] sets out field equations that describe a matter-free field. A German military officer, Karl Schwarzschild [2], shortly before he died, derived a solution of the field equations for a static gravitational field of spherical symmetry. Schwarzschild's solution is referred to as the Schwarzschild metric.

Einstein [3] showed that matter cannot be compacted below a critical radius defined by the Schwarzschild metric. Weller [4] shows that compacting matter below the critical radius to form a black hole results in a violation of the conservation of momentum and energy.

Why, then, do many believe that black holes exist in Einstein's relativistic space time? The belief appears to have arisen based, at least partly, on an incorrect description of the journey of a particle falling radially towards a hypothetical mass compacted below the critical radius. The description is incorrect in that the particle reaches and crosses the critical radius.

Herein are discussed five fallacies used in the description of the particle's journey. Preliminary to addressing the fallacies, it is shown why the particle will never reach the critical radius.

2 Mapping coordinate time t to local time τ

For a particle falling radially toward a hypothetical mass compacted below a critical radius, a mapping of the coordinate time t of a distant observer to a local time τ of the particle based on a radial distance r is shown in Fig. 1. The data shown in Fig. 1 can be obtained using the Schwarzschild metric.

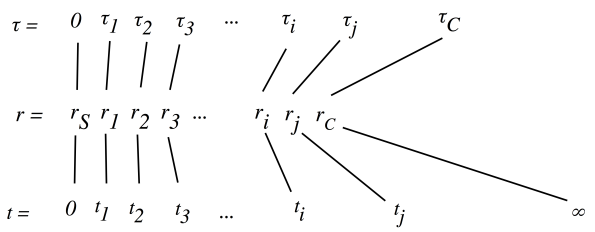


Fig. 1: For a particle falling radially, the Schwarschild Metric maps every value of the coordinate time t of a distant observer — where $0 \le t \le \infty$ — into a corresponding value of the local time τ of the particle — where $0 \le \tau \le \tau_C$.

Particularly, for a compact mass M with a Schwarzschild radius R, the Schwarzschild metric can be expressed using reference space coordinates (r, θ, ϕ), a coordinate time t and a local time τ (often referred to as proper time τ), i.e.,

$$c^2 d\tau^2 = c^2\left(1-\frac{R}{r}\right)dt^2 - \frac{dr^2}{(1-R/r)} - r^2 d\theta^2 - (r^2 sin^2\theta)d\phi^2. \quad (1)$$

Reference coordinates (r, θ, ϕ, t) are the space and time coordinates used by the distant observer to make measurements while the particle detects passage of time using local time coordinate τ. For a particle falling radially

$$d\theta = d\phi = 0, \quad (2)$$

so the Schwarzschild metric in (1) reduces to

$$c^2 d\tau^2 = c^2\left(1-\frac{R}{r}\right)dt^2 - \frac{dr^2}{(1-R/r)}, \quad (3)$$

which expresses a relationship between radial location r, local time τ and coordinate time t.

According to the relationship expressed by (3), for every radial location r_i reached from a starting location r_S, the coordinate time t_i to reach radial location r_i can be calculated using an integral

$$t_i = \int_{r_S}^{r_i} dt = \int_{r_S}^{r_i} f_1(r)dr, \quad (4)$$

where $f_1(r)$is a function of r derived from (3) [5, p. 667].

The local time τ_i required to reach the radial location r_i can be calculated using an integral

$$\tau_i = \int_{r_S}^{r_i} d\tau = \int_{r_S}^{r_i} f_2(r)dr, \quad (5)$$

where $f_2(r)$is a function of r derived from (3) [5, p. 663].

When the radial location r_i is set equal to a critical radius r_C, the integrand $f_1(r)$ for the integral in (4) and the integrand $f_2(r)$ for the integral in (5) are undefined; however, the integral in (5) converges while the integral in (4) does not. This

indicates that the critical radius r_C is reached in a finite local time τ_C but cannot be reached in finite Schwarzschild coordinate time.

The results of calculations using the integral of (4) and the integral of (5) are summarized in Fig. 1. As shown by Fig. 1, based on the integrals in (4) and (5), any value of coordinate time t, $0 \le t \le \infty$, can be mapped into a corresponding value for local time τ, $0 \le \tau \le \tau_C$ based on radial location r.

3 A pause to check correctness of Fig. 1

At this point the reader is encouraged to stop, look at Fig. 1, and perform an obviousness check to confirm why the data in Fig. 1 must be correct. The salient points are as follows:

- It takes infinite coordinate time (i.e., $t = \infty$) to reach the critical radius r_C;
- It takes finite local time τ_C to reach the critical radius r_C;
- Both local time τ and coordinate time t monotonically progress with decreasing r;
- To reach each radial location r_i will take a coordinate time t_i to complete and a local time τ_i to complete;
- Based on radial location r_i, a value of coordinate time t_i is mapped to a local time τ_i.

A reader who understands why Fig. 1 must be an accurate description of data derived from the Schwarzschild metric has already made a paradigm shift which if held to provides an intuitive foundation from which to understand the remainder of the paper. There is only one slight modification to Fig. 1 that is necessary to reveal why the critical radius can never be crossed. That is the subject of the next section.

4 Fig. 1 modified to take into account the finite duration of the compact mass

Fig. 1 depicts data from the Schwarzschild metric for a hypothetical compact mass that is presumed to exist forever in coordinate time. But what happens when the compact mass is replaced by an entity that more closely approximates reality in that it has a finite lifetime? For example, replace the compact mass with a theoretical black hole that has a finite lifetime. The result is shown in Fig. 2.

Because of Hawking radiation [6], it is estimated that a black hole will evaporate well within 10^{100} years. Therefore, added to Fig. 2 is finite coordinate time t_E which is the coordinate time required for a hypothetical black hole to completely evaporate [7]. Using the mapping shown in Fig. 1, it is possible to identify a radial location r_E — where $r_E > r_C$ — the particle will have reached simultaneous with the black hole evaporating at coordinate time t_E.

Fig. 2 shows a local time τ_E that represents the local time required for the particle to reach r_E. Local time τ_E corresponds with coordinate time t_E — the coordinate time required for a black hole to completely evaporate. Local time

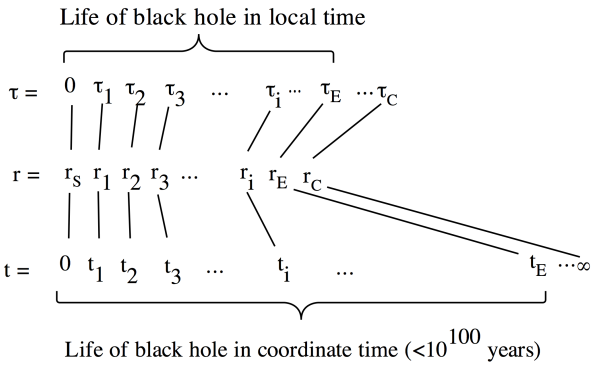


Fig. 2: According to the mapping of coordinate time to local time performed using the Schwarschild metric, the local time required to reach the critical radius of a black hole (τ_C) is longer than the life of the black hole (τ_E).

τ_C, as calculated by (5), represents the local time required for the particle to reach critical radius r_C. Because $\tau_E < \tau_C$, the particle will experience in local time τ that the black hole will evaporate before the critical radius can be reached.

5 The significance of Fig. 2

Fig. 2, based on the data from the Schwarzschild metric, shows a radially falling particle will never cross the critical radius of the compact mass regardless of what coordinate is used to measure the passage of time. For every radial location reached by the particle (i.e., $r_S \ge r \ge r_E$, there is a corresponding coordinate time t to reach the radial location and a corresponding local time τ to reach the radial location. The final destination of the particle is not dependent upon which measure of time is used to time the journey.

Fig. 2 presents a paradigm that is in conformance with the fundamental requirement of general relativity — and indeed a coherent universe — that there is a single reality with a logical sequence of events. The logical sequence of events does not vary based upon the reference frame from which observations are made.

Fig. 2 is meant to be an anchor from which can be shown how each of the five fallacies discussed below entices a departure from a coherent reality, where the logical sequence of events is consistent for every reference frame, into an incoherent reality where physical events differ based on reference frames from which observations are made.

In the following discussion of fallacies, evaporation of black holes is used as a convenient way to account for the finite lifetime of a hypothetical mass compacted below the critical radius. However, as should be clear from Fig. 2, a particle cannot cross the critical radius and therefore, as pointed out by [3], a mass will never compact below its critical radius. For the implication of this for collapsing stars, see the discussion of fallacy 4 below.

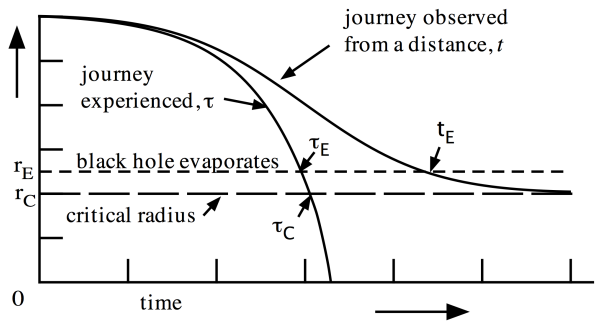


Fig. 3: Fig. 3 arranges the data shown in Fig. 2 in a different format. The trace extending to $r = 0$ incorrectly suggests that it is physically possible to cross the critical radius.

6 Fallacy 1: Showing a particle crosses the critical radius after evaporation of a black hole

For the journey of a particle to a black hole, elapsed time calculated using (4) and (5) is typically not represented as set out in Fig. 2, but rather as set out in Fig. 3 [5, p. 667].

Fig. 3, like Fig. 1 and Fig. 2, is a graphic representation of data obtained from (4) and (5). However, Fig. 3 qualifies as a fallacy because Fig. 3 includes extra data, not shown in Fig. 1 or Fig. 2., that incorrectly portrays the journey of the particle. Particularly, in Fig. 3, the trace representing local time τ extends beyond τ_C, the local time required to reach critical radius r_C.

Ordinary rules of mathematics cannot be used to generate the extra data for local time τ that occur after critical radius r_C is reached. This is because the integrand in (5) is undefined at r_C. Nevertheless, a novel "cycloid principle" [5, See pp. 663–664] has been used to generate this extra data.

But merely showing how the extra data can be mathematically generated does not overcome the logical sequencing problem introduced by adding the extra data to Fig. 3. The extra data suggests r_C can be reached and crossed in local time τ_C. However, this is impossible because as shown in Fig. 2, a black hole will evaporate in local time τ_E, so that critical radius r_C will cease to exist before it can be reached by the particle.

A horizontal line has been included in Fig. 3 to indicate where in Fig. 3 the evaporation of a black hole occurs. As shown by Fig. 3, evaporation of a black hole at radial location r_E, local time τ_E and coordinate time t_E logically occurs before reaching radial location r_C, local time τ_C and coordinate time $t = \infty$.

Fig. 3 should be corrected to show that a physical journey of a particle towards a black hole must end at radial location r_E — short of the critical radius r_C — when the black hole evaporates at local time τ_E and coordinate time t_E. The end of the journey occurs at r_E whether time is measured using coordinate time t or local time τ.

7 Fallacy 2: Declaring coordinates to be "pathological"

Fig. 3 suggests an impossible picture of physical reality. The particle cannot finally arrive at different destinations ($r = 0$ and $r = r_C$) merely based on the coordinate used to measure time.

As discussed in the last section, the logical sequence of events that occurs in all time frames, as out in Fig. 2, makes clear what is wrong with Fig. 3 and how it can be corrected. However, another competing explanation has been put forth.

The infinite coordinate time t required to reach the critical radius has been explained as the result of a "pathology" in the coordinates used to express the Schwarzschild metric. [5, See pp. 820-823].

Declaring coordinates to be pathological is a fallacy because it is a violation of general relativity at its most fundamental level. According to general relativity, all coordinates (reference frames) will observe the same reality. As Einstein [1, p. 117] made clear when setting out the basis for the theory of general relativity: "... all imaginable systems of coordinates, on principle, [are] equally suitable for the description of nature".

If general relativity is true, the events that occur during the journey of the particle occur in the same logical sequence irrespective of the coordinates used to observe the journey. Fig. 2 shows that the logical sequence of events that happens when time is measured using coordinate time t also happens in the same logical order when time is measured using local time τ. The next section shows that even when making observation from specially selected coordinates, the logical sequence of events does not differ from that shown in Fig. 2.

8 Fallacy 3: Use of specially selected coordinates

Fallacy 3 is an attempt to find coordinates that will show the particle can reach and cross the critical radius. The specially selected coordinates achieve this purpose based on a logical fallacy called begging the question in which the thing to be proved is assumed in a premise.

The thing to be proved is that a free falling particle can reach and cross the critical radius. The premise is that the specially selected coordinates can reach and cross the critical radius. When the specially selected coordinates are used as the reference coordinates in the Schwarzschild metric, and it is assumed the specially selected coordinates can cross the critical radius, it is possible to "show" the particle also can cross the critical radius.

But the premise is false. In the Schwarzschild metric, no reference frame can cross its critical radius because to do so would be a violation of the conservation of momentum and energy [4]. Below are considered two classes of specially selected coordinates:

- Coordinates that use the same reference frame as the free falling particle (e.g., the Novikov coordinates);

- Coordinates that use the reference frame of a radially traveling photon, (e.g., ingoing Eddington-Finkelstein coordinates and the Kruskal-Szekeres coordinates).

For each class of specially selected coordinates it is shown that their reference frame cannot cross a critical radius within the time it takes a black hole to evaporate.

Coordinates that use the reference frame of the free falling particle: Coordinates, such as the Novikov coordinates, that share a reference frame with the particle, also share the same time coordinate. Thus the local time coordinate τ measures the passage of time for both the local coordinates and the reference frame of the Novikov coordinates [5, p. 826].

The time required for a black hole to evaporate as measured by the time coordinate τ — which is the time coordinate for the reference frame shared by the Novikov coordinates shared and the local coordinates — has already been shown to be τ_E. See Fig. 2. As discussed above, $\tau_E < \tau_C$, indicating a black hole will evaporate at local time τ_E before the reference frame for the Novikov coordinates and the particle will be able to reach the critical radius at local time τ_C.

Coordinates that use the reference frame of a photon: The reference frame for ingoing Eddington-Finkelstein coordinates and the Kruskal-Szekeres coordinates is a radially traveling photon. [5, See pp. 826–832].

The coordinate time t for the photon to reach its critical radius can be very simply calculated from the Schwarzschild metric in (1). Because the photon is traveling radially, $d\theta = d\phi = 0$. Because local time for a photon does not progress, $d\tau = 0$. Therefore, the form of the Schwarzschild metric used to calculate values for coordinate time t is obtained by setting $d\theta = d\phi = d\tau = 0$ in (1) yielding

$$0 = c^2\left(1 - \frac{R}{r}\right)dt^2 - \frac{dr^2}{(1 - R/r)}. \qquad (6)$$

The integral in (4) can be used to calculate elapsed coordinate time t for the photon based on radial distance. Integrand $f_1(r)$ is obtained by rearranging the terms in (6), i.e.,

$$f_1(r) = \frac{dt}{dr} = \frac{1}{c(1 - R/r)}. \qquad (7)$$

When the photon reaches $r = R$, the integrand in (7) is undefined and the integral in (4) does not converge. Therefore the radially traveling photon will not reach R in finite coordinate time.

A black hole that evaporates in finite coordinate time t_E, will evaporate when the photon reaches a radial location r_L that is outside R. When the photon reaches radial location r_L at coordinate time t_E, the ingoing particle will be at radial location r_E, outside the critical radius r_C, as shown by Fig. 2.

In the reference frame of a photon, the black hole will evaporate when the photon reaches radial location r_L, before the photon reaches its critical radius R. As in all reference frames of the Schwarzschild metric, the reference frame of the photon is not able to reach the critical radius before the black hole evaporates.

9 Fallacy 4: Claiming the existence of surfaces trapped below a surface of last influence

Misner et al. [5, pp. 873–874] makes the argument that once the surface of a collapsing star crosses a critical radius, light reflecting from the surface remains trapped below the critical radius. This is a fallacy because the surface of a collapsing star will never cross the critical radius [3]. The very last particle on the surface to cross the critical radius can be approximately modeled by the radially falling particle of Fig. 2. From the perspective of the distant observer (coordinate time in Fig. 2), the collapsing star evaporates in finite time, before the infinite coordinate time required for the last particle on the surface to cross the critical radius.

From the perspective of a particle on the surface (local time in Fig. 2), the collapsing star evaporates very suddenly as the particle nears the critical radius. It is intriguing to imagine the experience of the particle as the surface of the collapsing star immediately disintegrates into radiation near the critical radius. Such an inferno of unimaginable proportions would tend to be masked from a distant observer by the extreme gravity near the critical radius. But as the surface burns away reducing the mass of the collapsing star — causing the critical radius to retreat farther below the surface of the collapsing star — a less time dilated view of the inferno might be released, perhaps providing an explanation for the sudden appearance of quasars.

Since the surface of a collapsing star cannot cross its critical radius in finite coordinate time t, Misner et al. [5, pp. 873–874] measures time from the reference frame for the ingoing Eddington-Finkelstein coordinates. As discussed in the prior section, use of ingoing Eddington-Finkelstein coordinates to prove the critical radius can be crossed just begs the question. The ingoing Eddington-Finkelstein coordinates will not cross the Schwarzschild metric of the collapsing star before the collapsing star evaporates. This should be especially clear for the example of a collapsing star since the surface, located outside its critical radius, will be an impenetrable barrier that will prevent any photon, serving as a reference frame for the ingoing Eddington-Finkelstein coordinates, from reaching its critical radius at R.

10 Fallacy 5: Claiming the infinite coordinate time to reach the critical radius is an optical illusion

It has been asserted that as measured by proper time, a free-falling traveler quickly reaches the critical radius. To the distant observer it appears to take an infinite amount of coordinate time to reach the critical radius as a result of an optical illusion caused by light propagation introducing a delay in communicating that the critical radius has been reached [5,

pp. 874–875]. Fallacy 5 is a departure from general relativity because in general relativity the difference between local time and coordinate time is not merely the result of delay introduced by light propagation. In the theory of general relativity, time progresses at different rates depending on the strength of the gravity field in which measurements are made.

Einstein [8, p. 106] explains: "we must use clocks of unlike constitution, for measuring time at places with differing gravitational potential." This principle of relativity is embodied in the Schwarzschild metric where gravity changes the rate at which time progresses [2]. For a precise description of how in the Schwarzschild metric gravity affects time based on the conservation of momentum and energy, see [4, Eq. 8].

Because fallacy 5 does not properly account for the effect gravity has on time, and is therefore not in accord with general relativity or the Schwarzschild metric, the results predicted by fallacy 5 do not agree with results calculated using the Schwarzschild metric. This is illustrated by a hypothetical in the following section.

11 A hypothetical illustrating the logical contradictions introduced by fallacy 5

According to fallacy 5, as measured by proper time, a radially falling traveler quickly reaches and crosses the critical radius of a black hole. The reality that the traveler quickly reaches the critical radius appears to the distant observer to take an infinite amount of time because of the propagation of light.

Fallacy 5's portrayal of reality is not consistent with calculations made using the Schwarzschild metric.

For example, put a reflector on the back of the traveler and have the distant observer periodically shine a light beam at the traveler. Use the Schwarzschild metric to calculate the radial location at which the faster moving light beam will overtake the slower moving traveler and reflect back to indicate the location of the traveler to the distant observer.

No matter how much of a head start the traveler has before the light is turned on (even trillions of years or longer, as measured using coordinate time), according to the Schwarzschild metric the light will always overtake the traveler before the critical radius is reached. The radial location at which the traveler is overtaken is the same whether local time or coordinate time is used to make the calculations, provided start time and overtake time for each light beam are measured with the same time coordinate. This result is inevitable based on the pattern of the data obtained from the Schwarzschild metric, as shown in Fig. 1.

As shown by Fig. 2, the distant observer can continue to shine light beams at the traveler until the distant observer observes the black hole evaporates. The feedback from the reflected light beams will tell the distant observer that the traveler remains outside the black hole as the black hole evaporates slowly in coordinate time, and quickly in local time. This contradicts the assertion of fallacy 5 that the traveler easily reaches and crosses the critical radius.

The distant observer does not even need to shine a light beam for this experiment as background radiation reflecting from the traveler provides exactly the same information.

Hawking radiation also provides the same information. While the distant observer sees the traveler outside the critical radius, the distant observer will also observe Hawking radiation from the evaporating black hole, which will first have to pass through the radial location of the traveler before reaching the distant observer. This indicates to the distant observer that the traveler will have experienced, before the distant observer, radiation emitted during the disintegration of the black hole. Further, the radiation passing by the traveler will continuously bring information to the distant observer about the location of the traveler confirming the information from the light beams. Each photon of radiation from the evaporating black hole that passes by the traveler is a progress report on the traveler's location that will confirm to the distant observer that the traveler had not yet passed through the critical radius when that photon of radiation passed the traveler. Such progress reports will continue until the black hole completely evaporates.

Light beams from the distant observer, background radiation and Hawking radiation will all intercept the traveler outside the critical radius — according to the Schwarzschild metric — regardless of the coordinates used to make measurements. This result contradicts the assertion of fallacy 5 that the critical radius is quickly crossed and only appears to the distant observer to take infinite time because of light propagation.

Submitted on November 16, 2010 / Accepted on December 15, 2010

References

1. Einstein A. The foundation of the general theory of relativity. The Principle of Relativity. Dover Publications, New York, 1923. pp. 111–164.
2. Schwarzschild K. On the gravitational field of a mass point according to Einstein's theory. Translated by S. Antoci, A. Loinger, 1999, arXiv:Physics/9905030.
3. Einstein A. On a stationary system with spherical symmetry consisting of many gravitating masses. *Annals of Mathematics*, 1939, v. 40, no. 4, 922–936.
4. Weller D. How black holes violate the conservation of energy. *Progress in Physics*, 2011, v. 1, 89–92.
5. Misner C., Thorne K., Wheeler J. Gravitation, W.H. Freeman & Co, New York, 1973.
6. Hawking S. W. Black Hole Explosions? *Nature*, 1974, v. 248, no. 5443, 30–31.
7. Vachaspati T., Stojkovic D., Krauss L. Observation of incipient black holes and the information loss problem. *Physical Review D*, 2007, v. 76, 024005.
8. Einstein A. On the Influence of gravitation on the propagation of light. The Principle of Relativity. Dover Publications, New York, 1923, pp. 99–108.

Lee Smolin Five Great Problems and Their Solution without Ontological Hypotheses

Gunn Quznetsov

Chelyabinsk State University, Chelyabinsk, Ural, Russia. E-mail: gunn@mail.ru, quznets@yahoo.com

Solutions of Lee Smolin Five Great Problems from his book *The Trouble with Physics: the Rise of String Theory, the Fall of a Science, and What Comes Next* are described. These solutions are obtained only from the properties of probability without any ontological hypotheses.

Introduction

In his book [1] Lee Smolin, professor of Perimeter Institute, Canada, has formulated the following five problems which he named Great Problems:

> *Problem 1: Combine general relativity and quantum theory into a single theory that claim to be the complete theory of nature.*
>
> *Problem 2: Resolve the problems in the foundations of quantum mechanics, either by making sense of the theory as it stands or by inventing a new theory that does make sense. ...*
>
> *Problem 3: Determine whether or not the various particles and forces can be unified in a theory that explain them all as manifestations of a single, fundamental entity. ...*
>
> *Problem 4: Explain how the values of of the free constants in the standard model of particle physics are chosen in nature. ...*
>
> *Problem 5: Explain dark matter and dark energy. Or if they don't exist, determine how and why gravity is modified on large scales. ...*

Solution

Let us consider the free Dirac Lagrangian:

$$\mathcal{L} := \psi^{\dagger}\left(\beta^{[k]}\partial_k + m\gamma^{[0]}\right)\psi. \qquad (1)$$

Here*

$$\beta^{[\nu]} := \begin{bmatrix} \sigma_\nu & 0_2 \\ 0_2 & -\sigma_\nu \end{bmatrix} \quad , \gamma^{[0]} := \begin{bmatrix} 0_2 & 1_2 \\ 1_2 & 0_2 \end{bmatrix}$$

where σ_1, σ_2, σ_3 are the Pauli matrices.

Such Lagrangian is not invariant [2] under the SU(2) transformation with the parameter α:

$$\psi^{\dagger}U^{\dagger}(\alpha)\left(\beta^{[k]}\partial_k + m_1\gamma^{[0]}\right)U(\alpha)\psi = \psi^{\dagger}\left(\beta^{[k]}\partial_k + (m\cos\alpha)\,\gamma^{[0]}\right)\psi,$$

the mass member is changed under this transformation.

Matrices $\beta^{[\nu]}$ and $\gamma^{[0]}$ are anticommutative. But it turns out that there exists a fifth matrix $\beta^{[4]}$ anticommuting with all these four matrices:

$$\beta^{[4]} := \mathrm{i}\begin{bmatrix} 0_2 & 1_2 \\ -1_2 & 0_2 \end{bmatrix}.$$

And the term with this matrix should be added to this Lagrangian mass term:

$$\underline{\mathcal{L}} := \psi^{\dagger}\left(\beta^{[k]}\partial_k + m_1\gamma^{[0]} + m_2\beta^{[4]}\right)\psi$$

where $\sqrt{m_1^2 + m_2^2} = m$.

Let $U(\alpha)$ be any SU(2)-matrix with parameter α and let $\mathbf{U}$ be the space in which $U(\alpha)$ acts. In such case $U(\alpha)$ divides the space $\mathbf{U}$ into two orthogonal subspaces $\mathbf{U}_o$ and $\mathbf{U}_x$ such that for every element ψ of $\mathbf{U}$ there exists an element ψ_o of $\mathbf{U}_o$ and an element ψ_x of $\mathbf{U}_x$ which fulfills the following conditions [3, 4]:

1.
$$\psi_o + \psi_x = \psi,$$

2.
$$\begin{aligned}&\psi_o^{\dagger}U^{\dagger}(\alpha)\left(\beta^{[k]}\partial_k + m_1\gamma^{[0]} + m_2\beta^{[4]}\right)U(\alpha)\psi_o = \\ =\ &\psi_o^{\dagger}(\beta^{[k]}\partial_k + (m_1\cos\alpha - m_2\sin\alpha)\,\gamma^{[0]} + \\ &+ (m_2\cos\alpha + m_1\sin\alpha)\,\beta^{[4]})\psi_o,\end{aligned} \qquad (2)$$

3.
$$\begin{aligned}&\psi_x^{\dagger}U^{\dagger}(\alpha)\left(\beta^{[k]}\partial_k + m_1\gamma^{[0]} + m_2\beta^{[4]}\right)U(\alpha)\psi_x = \\ =\ &\psi_x^{\dagger}(\beta^{[k]}\partial_k + (m_1\cos\alpha + m_2\sin\alpha)\,\gamma^{[0]} + \\ &+ (m_2\cos\alpha - m_1\sin\alpha)\,\beta^{[4]})\psi_x.\end{aligned} \qquad (3)$$

In either case, m does not change.

I call these five ($\beta := \left\{\beta^{[\nu]}, \beta^{[4]}, \gamma^{[0]}\right\}$) anticommuting matrices *Clifford pentad*. Any sixth matrix does not anticommute with all these five.

There exist only six Clifford pentads (for instance, [5, 6]): I call one of them (the pentad β) *the light pentad*, three (ζ, η, θ) — *the chromatic pentads*, and two ($\underline{\Delta}$, $\underline{\Gamma}$) — *the gustatory pentads*.

*$0_2 := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, 1_2 := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \beta^{[0]} := -1_4 := -\begin{bmatrix} 1_2 & 0_2 \\ 0_2 & 1_2 \end{bmatrix}$, $k \in \{0, 1, 2, 3\}$, $\nu \in \{1, 2, 3\}$.

The light pentad contains three diagonal matrices ($\beta^{[\nu]}$) corresponding to the coordinates of 3-dimensional space, and two antidiagonal matrices ($\beta^{[4]}$, $\gamma^{[0]}$) relevant to mass terms (2,3) — one for the lepton state and the other for the neutrino state of this lepton.

Each chromatic pentad also contains three diagonal matrices corresponding to three coordinates and two antidiagonal mass matrices - one for top quark state and the other — for bottom quark state.

Each gustatory pentad contains a single diagonal coordinate matrix and two pairs of antidiagonal mass matrices [6] — these pentads are not needed yet.

Let* $\langle \rho_A \mathrm{c},\, j_{A,\nu} \rangle$ be a 1+3-vector of probability density of a pointlike event A.

For any A the set of four equations with an unknown complex 4×1 matrix function $\varphi(x_k)$

$$\left\{ \begin{array}{c} \rho_A = \varphi^\dagger \varphi, \\ \dfrac{j_{A,\nu}}{\mathrm{c}} = -\varphi^\dagger \beta^{[\nu]} \varphi \end{array} \right|$$

has solution [3].

If† $\rho_{\mathfrak{A}}(x_k) = 0$ for all x_k such that $|x_k| > (\pi \mathrm{c}/\mathrm{h})$ then φ obeys the following equation [10]:

$$\begin{aligned} &\Big(-\left(\partial_0 + \mathrm{i}\Theta_0 + \mathrm{i}\Upsilon_0\gamma^{[5]}\right) + \beta^{[\nu]}\left(\partial_\nu + \mathrm{i}\Theta_\nu + \mathrm{i}\Upsilon_\nu\gamma^{[5]}\right) + \\ &\qquad + 2\left(\mathrm{i}M_0\gamma^{[0]} + \mathrm{i}M_4\beta^{[4]}\right)\Big)\varphi + \\ &+\Big(-\left(\partial_0 + \mathrm{i}\Theta_0 + \mathrm{i}\Upsilon_0\gamma^{[5]}\right) - \zeta^{[\nu]}\left(\partial_\nu + \mathrm{i}\Theta_\nu + \mathrm{i}\Upsilon_\nu\gamma^{[5]}\right) + \\ &\qquad + 2\left(-\mathrm{i}M_{\zeta,0}\gamma^{[0]}_\zeta + \mathrm{i}M_{\zeta,4}\zeta^{[4]}\right)\Big)\varphi + \\ &+\Big(\left(\partial_0 + \mathrm{i}\Theta_0 + \mathrm{i}\Upsilon_0\gamma^{[5]}\right) - \eta^{[\nu]}\left(\partial_\nu + \mathrm{i}\Theta_\nu + \mathrm{i}\Upsilon_\nu\gamma^{[5]}\right) + \\ &\qquad + 2\left(-\mathrm{i}M_{\eta,0}\gamma^{[0]}_\eta - \mathrm{i}M_{\eta,4}\eta^{[4]}\right)\Big)\varphi + \\ &+\Big(-\left(\partial_0 + \mathrm{i}\Theta_0 + \mathrm{i}\Upsilon_0\gamma^{[5]}\right) - \theta^{[\nu]}\left(\partial_\nu + \mathrm{i}\Theta_\nu + \mathrm{i}\Upsilon_\nu\gamma^{[5]}\right) + \\ &\qquad + 2\left(\mathrm{i}M_{\theta,0}\gamma^{[0]}_\theta + \mathrm{i}M_{\theta,4}\theta^{[4]}\right)\Big)\varphi = \\ &= \ 0 \end{aligned}$$

with real
$\Theta_k(x_k)$, $\Upsilon_k(x_k)$, $M_0(x_k)$, $M_4(x_k)$, $M_{\zeta,0}(x_k)$, $M_{\zeta,4}(x_k)$, $M_{\eta,0}(x_k)$, $M_{\eta,4}(x_k)$, $M_{\theta,0}(x_k)$, $M_{\theta,4}(x_k)$
and with

$$\gamma^{[5]} := \left[\begin{array}{cc} 1_2 & 0_2 \\ 0_2 & -1_2 \end{array} \right].$$

*c = 299792458.

†h := $6.6260755 \cdot 10^{-34}$

The first summand of this equation contains elements of the light pentad only. And the rest summands contain elements of the chromatic pentads only.

This equation can be rewritten in the following way:

$$\begin{aligned} &\beta^{[k]}\left(-\mathrm{i}\partial_k + \Theta_k + \Upsilon_k\gamma^{[5]}\right)\varphi + \\ &+(M_0\gamma^{[0]} + M_4\beta^{[4]} - \\ &-M_{\zeta,0}\gamma^{[0]}_\zeta + M_{\zeta,4}\zeta^{[4]} - \\ &-M_{\eta,0}\gamma^{[0]}_\eta - M_{\eta,4}\eta^{[4]} + \\ &+M_{\theta,0}\gamma^{[0]}_\theta + M_{\theta,4}\theta^{[4]})\varphi = \\ = \ &0 \end{aligned} \tag{4}$$

because

$$\zeta^{[\nu]} + \eta^{[\nu]} + \theta^{[\nu]} = -\beta^{[\nu]}.$$

This equation is a generalization of the Dirac's equation with gauge fields $\Theta_k(x_k)$ and $\Upsilon_k(x_k)$ and with eight mass members. The mass members with elements of the light pentad (M_0 and M_4) conform to neutrino and its lepton states. And six mass members with elements of the chromatic pentads conform to three pairs (up and down) of chromatic states (red, green, blue).

Let this equation not contains the chromatic mass nembers:

$$\left(\beta^{[k]}\left(-\mathrm{i}\partial_k + \Theta_k + \Upsilon_k\gamma^{[5]}\right) + M_0\gamma^{[0]} + M_4\beta^{[4]}\right)\varphi = 0. \tag{5}$$

If function φ is a solution of this equation then φ represents the sum of functions $\varphi_{n,s}$ which satisfy the following conditions [3, 62–71]:

n and s are integers;

each of these functions obeys its equation of the following form:

$$\left(\beta^{[k]}\left(\mathrm{i}\partial_k - \Theta_0 - \Upsilon_0\gamma^{[5]}\right) - \frac{\mathrm{h}}{\mathrm{c}}\left(\gamma^{[0]}n + \beta^{[4]}s\right)\right)\varphi_{n,s} = 0; \tag{6}$$

for each point x_k of space-time: or this point is empty (for all n and s: $\varphi_{n,s}(x_k) = 0$), or in this point is placed a single function (for x_k there exist integers n_0 and s_0 such that $\varphi_{n_0,s_0}(x_k) \neq 0$ and if $n \neq n_0$ and/or $s \neq s_0$ then $\varphi_{n,s}(x_k) = 0$).

In this case if $m := \sqrt{n^2 + s^2}$ then m is a natural number. But under the SU(2)-transformation with parameter α (2, 3):
$m \to \left((n\cos\alpha - s\sin\alpha)^2 + (s\cos\alpha + n\sin\alpha)^2\right)^{0.5}$,
$(n\cos\alpha - s\sin\alpha)$ and $(s\cos\alpha + n\sin\alpha)$ must be integers too. But it is impossible.

But for arbitrarily high accuracy in distant areas of the natural scale there exist such numbers m that for any α some natural numbers n' and s' exist which obey the following conditions: $n' \approx (n\cos\alpha - s\sin\alpha)$ and $s' \approx (s\cos\alpha + n\sin\alpha)$. These numbers m are separated by long intervals and determine the mass spectrum of the generations of elementary particles. Apparently, this is the way to solve Problem 4 because

the masses are one of the most important constants of particle physics.

The Dirac's equation for leptons with gauge members which are similar to electroweak fields is obtained [4, p. 333–336] from equations (5, 6). Such equation is invariant under electroweak transformations. And here the fields W and Z obey the Klein-Gordon type equation with nonzero mass.

If the equation (4) does not contain lepton's and neutrino's mass terms then the Dirac's equation with gauge members which are similar to eight gluon's fields is obtained. And oscillations of the chromatic states of this equation bend space-time. This bend gives rise to the effects of redshift, confinement and asymptotic freedom, and Newtonian gravity turns out to be a continuation of subnucleonic forces [10]. And it turns out that these oscillations bend space-time so that at large distances the space expands with acceleration according to Hubble's law [7]. And these oscillations bend space-time so that here appears the discrepancy between the quantity of the luminous matter in the space structures and the traditional picture of gravitational interaction of stars in these structures. Such curvature explains this discrepancy without the Dark Matter hypothesis [8] (Problem 5).

Consequently, the theory of gravitation is a continuation of quantum theory (Problem 1 and Problem 3).

Thus, concepts and statements of Quantum Theory are concepts and statements of the probability of pointlike events and their ensembles.

Elementary physical particle in vacuum behaves like these probabilities. For example, according to doubleslit experiment [9], if a partition with two slits is placed between a source of elementary particles and a detecting screen in vacuum then interference occurs. But if this system will be put in a cloud chamber, then trajectory of a particle will be clearly marked with drops of condensate and any interference will disappear. It seems that a physical particle exists only in the instants of time when some events happen to it. And in the other instants of time the particle does not exist, but the probability of some event to happen to this particle remains.

Thus, if no event occurs between an event of creation of a particle and an event of detection of it, then the particle does not exist in this period of time. There exists only the probability of detection of this particle at some point. But this probability, as we have seen, obeys the equations of quantum theory and we get the interference. But in a cloud chamber events of condensation form a chain meaning the trajectory of this particle. In this case the interference disappears. But this trajectory is not continuous — each point of this line has an adjacent point. And the effect of movement of this particle arises from the fact that a wave of probability propagates between these points.

Consequently, the elementary physical particle represents an ensemble of pointlike events associated with probabilities. And charge, mass, energy, momentum, spins, etc. represent parameters of distribution of these probabilities. It explains all paradoxes of quantum physics. Schrödinger's cat lives easily without any superposition of states until the microevent awaited by everyone occurs. And the wave function disappears without any collapse in the moment when event probability disappears after the event occurs.

Hence, entanglement concerns not particles but probabilities. That is when the event of the measuring of spin of Alice's electron occurs then probability for these entangled electrons is changed instantly in the whole space. Therefore, nonlocality acts for probabilities, not for particles. But probabilities can not transmit any information (Problem 2).

Conclusion

Therefore, Lee Smolin's Five Great Problems do have solution only using the properties of probabilities. These solutions do not require any dubious ontological hypotheses such as superstrings, spin networks, etc.

Submitted on December 15, 2010 / Accepted on December 16, 2010

References

1. Smolin L. The trouble with physics: the rise of string theory, the fall of a science, and what comes next. Houghton Mifflin, Boston, 2006.
2. Kane G. Modern Elementary Particle Physics. Addison-Wesley Publ. Comp., 1993, p. 93.
3. Quznetsov G. Probabilistic Treatment of Gauge Theories, in series *Contemporary Fundamental Physics*, ed. Dvoeglazov V., Nova Sci. Publ., N.Y., 2007.
4. Quznetsov G. It is not Higgs. *Prespacetime Journal*, 2010, v. 1, 314–343.
5. Madelung E. Die Mathematischen Hilfsmittel des Physikers. Springer Verlag, Berlin, Gottingen, Heidelberg, 1957, p. 12.
6. Quznetsov G. Logical Foundation of Theoretical Physics. Nova Sci. Publ., N.Y., 2006.
7. Quznetsov G. A. Oscillations of the Chromatic States and Accelerated Expansion of the Universe. *Progress in Physics*, 2010, v. 2, 64–65.
8. Quznetsov G. Dark Matter and Dark Energy are Mirage. *Prespacetime Journal*, October 2010, v. 1, Issue 8, 1241–1248, arXiv: 1004.4496 2.
9. Quznetsov G. Double-Slit Experiment and Quantum Theory Event-Probability Interpretation, arXiv: 1002.3425.
10. Quznetsov G. A. 4X1-Marix Functions and Dirac's Equation. *Progress in Physics*, 2009, v. 2, 96–106.

On the Failure of Particle Dark Matter Experiments to Yield Positive Results

Joseph F. Messina
Topical Group in Gravitation, American Physical Society,
P.O. Box 130520, The Woodlands, TX 77393, USA.
Email: jfmessina77@yahoo.com

It is argued that the failure of *particle* dark matter experiments to verify its existence may be attributable to a *non-Planckian* "action", which renders dark matter's behavior contradictory to the consequences of quantum mechanics as it applies to luminous matter. It is pointed out that such a possibility cannot be convincingly dismissed in the absence of a physical law that prohibits an elementary "action" smaller than Planck's. It is further noted that no purely dark matter measurement of Planck's constant exists. Finally, the possibility of a non-Planckian cold dark matter particle is explored, and found to be consistent with recent astronomical observations.

The search for dark matter (DM) remains one of the most vexing of the unresolved problems of contemporary physics. While the existence of DM is no longer in dispute, its composition is a matter of lively debate. A variety of subatomic particles with exotic properties have been proposed as possible candidates. However, as is well known by now, after more than three decades of experimentation, and considerable expenditure, none have yet been detected. If the past is any guide, such negative results often force us to radically reexamine some of the basic tenets underlying physical concepts. It is the purpose of this paper to propose a plausible, experimentally verifiable, explanation for the persistent failure of *particle* DM experiments to yield positive results.

Since DM's existence is inferred solely from its gravitational effects, and its nature is otherwise unknown, one cannot rule-out the possibility that DM's behavior may be contradictory to the consequences of quantum mechanics as it applies to luminous matter (LM), which is particularly troubling since it necessarily brings into question the applicability of Planck's constant as a viable "action" in this *non-luminous* domain. It is important to point out that *no* purely DM measurement of Planck's constant exists. Indeed, all that we know about Planck's constant is based on electromagnetic and strong interaction experiments, whose particles and fields account for only 4.6% of the mass-energy density of the observable universe, which pales when compared to the 23.3% attributable to DM.

While it is true that very little is known about DM, some progress has been made on the astronomical front. Recent observations have revealed important new clues regarding its behavior. Particularly important, an analysis of cosmic microwave background observables has provided conclusive evidence that DM is made up of slow-moving particles [1], a development that has firmly established the cold DM paradigm as the centerpiece of the standard cosmology. Equally revealing, large aggregates of DM have been observed passing right through each other without colliding [2–3], which is clearly significant since it essentially rules out the idea that particles of DM can somehow interact and collide with each other. Taken together these astronomical findings are suggestive of a non-relativistic, non-interacting, particle whose coherent mode of behavior is a characteristic of *classical light*. Clearly, for such a particle, the condition of quantization can only become a physical possibility if its "action" is considerably *smaller* than Planck's.

Upon reflection one comes to the realization that such a possibility can be accommodated in the context of the framework of quantum mechanics, whose formalism allows for *two* immutable "actions". Namely, Planck's familiar constant, h, which has been shown experimentally to play a crucial role in the microphysical realm, and the more diminutive, less familiar "action" e^2/c where e is the elementary charge, and c is the velocity of light in a vacuum (denoted by the symbol j for simplicity of presentation). While this *non-Planckian* constant appears to have no discernible role in our luminous world, it is, nevertheless, clearly of interest since it may be sufficiently *smaller* than Planck's constant to account for DM's astronomical behavior; a possibility that cannot be convincingly dismissed in the absence of a physical law that prohibits an *elementary* "action" smaller than Planck's.

Whether or not we know DM's nature, the undisputed fact remains that *all* elementary particles exhibit wavelike properties. Hence, if DM's behavior is orchestrated by this *non-Planckian* "action" it should be possible to describe such particle waves quantum mechanically. In order to facilitate matters we shall assume that DM's non-Planckian *particle/wave* properties are consistent with both the Einstein relation for the total energy of a particle, in the form

$$E = jf = mc^2 = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}} \tag{1}$$

and the de Broglie relation for the momentum

$$p = \frac{j}{\lambda} = mv = \frac{m_0 v}{\sqrt{1 - v^2/c^2}}\,, \tag{2}$$

where $j = 7.6956 \times 10^{-30}$ erg s is the conjectured DM "action" quantum, which may be compared with the Planck constant,

h, found in our luminous world (i.e., 6.6260×10^{-27} erg s). Now, since the relation between energy and momentum in *classical* mechanics is simply

$$E = \frac{1}{2m}\, p^2 \tag{3}$$

we can replace E and p with the differential operators

$$E = i\,\frac{j}{2\pi}\,\frac{\partial}{\partial t} \tag{4}$$

and

$$p = -\,i\,\frac{j}{2\pi}\,\frac{\partial}{\partial x} \tag{5}$$

and operate with the result on the wave function $\psi(x,t)$ that represents the de Broglie wave. We then obtain

$$i\,\frac{j}{2\pi}\,\frac{\partial\psi}{\partial t} = -\frac{(j/2\pi)^2}{2m}\,\frac{\partial^2\psi}{\partial x^2}\,, \tag{6}$$

which is Schrödinger's general wave equation for a non-relativistic *free* particle. Its solution describes a *non-Planckian* particle that is the quantum mechanical analog of a non-interacting *classical* particle that is moving in the x direction with constant velocity; a result that closely mirrors DM's elusive behavior, and can be simply explained in the context of this generalization. That is, the *classical* concept of two particles exerting a force on each other corresponds to the quantum mechanical concept that the de Broglie wave of one particle influences the de Broglie wave of another particle. However, this is only possible if the de Broglie wave propagates *non-linearly,* in sharp contrast with Schrödinger's general wave equation for which the propagation of waves is described by a *linear* differential equation. Hence the presence of one wave *does not* affect the behavior of another wave, allowing them to pass right through each other without colliding, which is consistent with the results of the aforementioned astronomical observations [2–3].

If it exists, this non-Planckian particle would easily have eluded detection because of the diminutive magnitude of the non-Planckian "action". More succinctly, the closer one comes to the *classical* limit the *less* pronounced are the quantum effects. As a result, its behavior is expected to be *more* wave-like than particlelike, which is consistent with the observed coherent mode of behavior of large aggregates of DM [2–3]. Clearly, the detection of this non-Planckian particle in a terrestrial laboratory setting will, almost certainly, require the use of a wholly different set of experimental tools than those presently employed in conventional DM experiments, which are, after all, specifically designed to detect *particle* interactions.

While, as has been shown, DM's behavior in the astronomical arena can be satisfactorily accounted for quantum mechanically, in terms of this non-Planckian "action", the detailed implications remain to be worked out. Nevertheless, the introduction of this *non-Planckian cold DM particle* in the context of quantum mechanics, provides a fundamentally plausible means of explaining the failure of *conventional* experiments to provide conclusive evidence for the *particle* nature of DM. After these many decades of null experimental results, the time has come to acknowledge the possibility that DM's behavior may be orchestrated by a richer variety of fundamentally different mechanisms than previously recognized.

Appendix

I have taken note of the fact that if the reader is to grapple with some of the concepts generated by this paper, it would be advisable to ascribe an appropriate name to this non-Planckian particle. Clearly, the basic aspect that one should be mindful of is this particle's indispensable role in enabling the *warping* of spacetime sufficiently enough to cradle whole galaxies. Hence, I believe *"warpton"* would be the name of choice.

It is hoped that the experimental community can be sufficiently motivated to make a determined search for this provocative particle.

Submitted on December 14, 2010 / Accepted on December 15, 2010.

References

1. Lewis A.D., Buote D.A., Stocke J.T. Chandra Observations of A2029: The Dark Matter Profile Down to below $0.01r_{\rm vir}$ in an Unusually Relaxed Cluster. *The Astrophysical Journal*, 2003, v. 586, 135–142.
2. Clowe D., Bradac M., Gonzalez A.H., Markevitch M., Randall S.W., Jones C., Zaritsky D. A Direct Empirical Proof of the Existence of Dark Matter. *The Astrophysical Journal*, 2006, v. 648(2), L109–L113.
3. Natarajan P., Kneib J.-P. Smail I., Ellis R. Quantifying Substructure Using Galaxy-Galaxy Lensing in Distant Clusters. arXiv: astro-ph/0411426.

Application of the Model of Oscillations in a Chain System to the Solar System

Andreas Ries and Marcus Vinicius Lia Fook
Universidade Federal de Campina Grande, Unidade Acadêmica de Engenharia de Materiais, Rua Aprígio Veloso 882, 58429-140 Campina Grande — PB, Brazil

E-mail: andreasries@yahoo.com

A numerical analysis revealed that masses, radii, distances from the sun, orbital periods and rotation periods of celestial bodies can be expressed on the logarithmic scale though a systematic set of numbers: $4e$, $2e$, e, $\frac{e}{2}$, $\frac{e}{4}$, $\frac{e}{8}$ and $\frac{e}{16}$. We analyzed these data with a fractal scaling model originally published by Müller in this journal, interpreting physical quantities as proton resonances. The data were expressed in continued fraction form, where all numerators are Euler's number. From these continued fractions, we explain the volcanic activity on Venus, the absence of infrared emission of Uranus and why Jupiter and Saturn emit more infrared radiation than they receive as total radiation energy from the Sun. We also claim that the Kuiper cliff was not caused by a still unknown planet. It can be understood why some planets have an atmosphere and others not, as well as why the ice on dwarf planet Ceres does not evaporate into space through solar radiation. The results also suggest that Jupiter and Saturn have the principal function to capture asteroids and comets, thus protecting the Earth, a fact which is well-reflected in the high number of their irregular satellites.

1 Introduction

Recently in three papers of this journal, Müller [1–3] suggested a chain of similar harmonic oscillators as a general model to describe physical quantities as proton resonance oscillation modes. In this model, the spectrum of eigenfrequencies of a chain system of many proton harmonic oscillators is given by a continuous fraction equation [2]:

$$f = f_p \exp S \tag{1}$$

where f is any natural oscillation frequency of the chain system, f_p the oscillation frequency of one proton and S the continued fraction corresponding to f. S was suggested to be in the canonical form with all partial numerators equal 1 and the partial denominators are positive or negative integer values.

$$S = n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \ldots}}} \tag{2}$$

Particularly interesting properties arise when the nominator equals 2 and all denominators are divisible by 3. Such fractions divide the logarithmic scale in allowed values and empty gaps, i.e. ranges of numbers which cannot be expressed with this type of continued fractions. He showed that these continued fractions generate a self-similar and discrete spectrum of eigenvalues [1], that is also logarithmically invariant. Maximum spectral density areas arise when the free link n_0 and the partial denominators n_i are divisible by 3.

In a previous article [5] we slightly modified this model, substituting all nominators by Euler's number. In that way we confirmed again that elementary particles are proton resonance states, since most masses were found to be located close to spectral nodes and definitively not random.

In this article we investigated various solar system data, such as masses, sizes and distances from the Sun, rotation and orbital periods of celestial bodies on the logarithmic scale. We showed that continued fractions with Euler's number as nominator are adequate to describe the solar system. From these continued fractions we derived claims regarding specific properties of planets. It became evident, that the solar system possesses a hidden fractal structure.

2 Data sources and computational details

All solar system data, such as distances, masses, radii, orbital and rotation periods of celestial bodies, were taken from the NASA web-site. The km was converted into the astronomical unit via 1 AU = 149, 597, 870.7 km. The mean distance of an object from the central body is understood as $\frac{1}{2}(Aphelion - Periphelion)$. Numerical values of continued fractions were always calculated using the the Lenz algorithm as indicated in reference [4].

3 Results and discussion

3.1 Standard numerical analysis

Before doing any numerical analysis, one always has to be aware of the fact that the numerical value of a quantity depends on the physical unit. In this particular analysis we decided to choose practical units which were made exclusively by nature. Such units are the astronomical unit (AU) for lengths, the earth mass for planetary masses, as well as the year and the day for orbit and rotation periods. As can be seen, this particular choice leads to quite interesting regularities.

In a previous article [5], we had already done a similar analysis of elementary particle masses on the logarithmic

scale and detected a set of systematic mass gaps: $2e$, e, $\frac{e}{2}$, $\frac{e}{4}$, $\frac{e}{8}$ and $\frac{e}{16}$. Therefore, our numerical analysis was focused on these numbers and in a similar way, we detected this set of expressions again.

When looking from the Earth in direction away from the Sun, it can be noted that there are two principal zones, where mass accumulation into heavy planets seems to be forbidden. The existing mass is scattered in the form of asteroids and large bodies cannot become more than dwarf planets. The first such zone is the so-called Asteroid belt, located between Mars and Jupiter. Its population has already been well investigated, especially to confirm the orbital resonance effects manifesting in the Kirkwood gaps. Most asteroids have semi-major axes between 2.1 and 3.5 AU.

The second scattered-mass zone is the Kuiper belt, located from the orbit of Neptune (30 AU) to 55 AU distance from the Sun.

The Oort cloud is also such a scattered-mass zone. Due to its giant distance from the center of the solar system, there is no well-confirmed lower and upper limit, so we did not include it into the numerical analysis.

Table 1: Mean distances of celestial bodies (d) from the Sun expressed through e on the logarithmic scale and absolute values of corresponding numerical errors.

Object d [AU] ln(d)	Expression	Numerical error
Mercury 0.3871044 -0.9491	$-\left(\frac{e}{4}+\frac{e}{8}\right)$	0.0703
Venus 0.723339 -0.3239	$-\frac{e}{8}$	0.0159
Earth 0.9999808 0.0000	$0e$	0.0000
Mars 1.523585 0.4211	$\frac{e}{8}$	0.0812
Ceres 2.7663 1.0175	$\frac{e}{4}+\frac{e}{8}$	0.0019
Jupiter 5.204419 1.6495	$\frac{e}{2}+\frac{e}{8}$	0.0494
Saturn 9.582516 2.2599	$\frac{e}{2}+\frac{e}{4}+\frac{e}{16}$	0.0513
Uranus 19.201209 2.9550	$e+\frac{e}{16}$	0.0668
Neptune 30.04762 3.4028	$e+\frac{e}{4}$	0.0049
Pluto 39.486178 3.6758	$e+\frac{e}{4}+\frac{e}{8}$	0.0618

Table 2: Equatorial radii (r) of celestial bodies expressed through e on the logarithmic scale and absolute values of corresponding numerical errors.

Object r [AU] ln(r)	Expression	Numerical error
Mercury 1.6308×10^{-5} -11.0238	$-\left(4e+\frac{e}{16}\right)$	0.0192
Venus 4.0454×10^{-5} -10.1154	$-\left(2e+e+\frac{e}{2}+\frac{e}{4}\right)$	0.0782
Earth 4.2635×10^{-5} -10.0628	$-\left(2e+e+\frac{e}{2}+\frac{e}{8}+\frac{e}{16}\right)$	0.0391
Mars 2.2708×10^{-5} -10.6928	$-\left(2e+e+\frac{e}{2}+\frac{e}{4}+\frac{e}{8}+\frac{e}{16}\right)$	0.0104
Ceres 3.2574×10^{-6} -12.6346	$-\left(4e+e+\frac{e}{2}+\frac{e}{8}\right)$	0.0625
Jupiter 4.7789×10^{-4} -7.6461	$-\left(2e+\frac{e}{2}+\frac{e}{4}+\frac{e}{16}\right)$	0.0010
Saturn 4.0287×10^{-4} -7.8169	$-\left(2e+\frac{e}{2}+\frac{e}{4}+\frac{e}{8}\right)$	0.0018
Uranus 1.709×10^{-4} -8.6747	$-\left(2e+e+\frac{e}{8}+\frac{e}{16}\right)$	0.0102
Neptune 1.6554×10^{-4} -8.7063	$-\left(2e+e+\frac{e}{8}+\frac{e}{16}\right)$	0.0418
Pluto 7.6940×10^{-6} -11.7751	$-\left(4e+\frac{e}{4}+\frac{e}{16}\right)$	0.0525
Sun 4.649×10^{-3} -5.3817	$-2e$	0.0549

Table 3: Sidereal orbital periods (T) of celestial bodies expressed through e on the logarithmic scale and absolute values of corresponding numerical errors.

Object T [y] ln(T)	Expression	Numerical error
Mercury 0.2408467 -1.4236	$-\frac{e}{2}$	0.0645
Venus 0.61519726 -0.4858	$-\left(\frac{e}{8}+\frac{e}{16}\right)$	0.0239
Earth 1.0000174 0.0000	$0e$	0.0000
Mars 1.8808476 0.6317	$\frac{e}{4}$	0.0479
Ceres 4.60 1.5261	$\frac{e}{2}+\frac{e}{16}$	0.0029
Jupiter 11.862615 2.4734	$\frac{e}{2}+\frac{e}{4}+\frac{e}{8}+\frac{e}{16}$	0.0750
Saturn 29.447498 3.3826	$e+\frac{e}{4}$	0.0153
Uranus 84.016846 4.4310	$e+\frac{e}{2}+\frac{e}{8}$	0.0138
Neptune 164.79132 5.1047	$e+\frac{e}{2}+\frac{e}{4}+\frac{e}{8}$	0.0079
Pluto 247.92065 5.5131	$2e$	0.0765

Table 4: Sidereal rotation periods (T) of celestial bodies (retrograde rotation ignored) expressed through e on the logarithmic scale and absolute values of corresponding numerical errors.

Object T [d] ln(T)	Expression	Numerical error
Mercury 58.6462 4.0715	$e+\frac{e}{2}$	0.0059
Venus 243.018 5.4931	$2e$	0.0565
Earth 0.99726968 -0.0027	$0e$	0.0027
Mars 1.02595676 0.0256	$0e$	0.0256
Ceres 0.3781 -0.9726	$-\left(\frac{e}{4}+\frac{e}{8}\right)$	0.0468
Jupiter 0.41354 -0.8830	$-\left(\frac{e}{4}+\frac{e}{16}\right)$	0.0335
Saturn 0.44401 -0.8119	$-\left(\frac{e}{4}+\frac{e}{16}\right)$	0.0376
Uranus 0.71833 -0.3308	$-\frac{e}{8}$	0.0090
Neptune 0.67125 -0.3986	$-\frac{e}{4}$	0.0588
Pluto 6.3872 1.8543	$\frac{e}{2}+\frac{e}{8}+\frac{e}{16}$	0.0145
Sun 25.05 3.2209	$e+\frac{e}{8}+\frac{e}{16}$	0.0071

It can be seen that the distance between Ceres (the largest Asteroid belt object) and Pluto (the largest Kuiper belt object) matches Euler's number quite accurately. Table 1 summarizes the mean distances of the most important celestial bodies from the Sun together with the corresponding natural logarithms. It was found that all logarithms can be expressed as a sum of $2e$, e, $\frac{e}{2}$, $\frac{e}{4}$, $\frac{e}{8}$ and $\frac{e}{16}$. Most distances could even expressed as multiples of $\frac{e}{8}$ since they do not contain the summand $\frac{e}{16}$. The numerical errors on the logarithmic scale are significantly lower than $\frac{e}{16}$.

Analogously, we expressed the equatorial radii, sidereal orbital periods, sidereal rotation periods and masses of celestial bodies on the logarithmic number line (see Tables 2–5). In very few cases it was necessary to introduce $4e$ into the set of summands.

From these results we conclude that all these numerical values of planetary data are definitively not a set of random numbers. The repeatedly occurring summands strongly support the idea of a self-similar, fractal structure as Müller already claimed in reference [2].

In the present form, these results are obtained only when considering nature-made units, which underlines their importance.

Table 5: Masses (m) of celestial bodies, rescaled by earth mass and expressed through e on the logarithmic scale and absolute values of corresponding numerical errors.

Object m [$\times 10^{24}$ kg] $\ln(\frac{m}{m_{Earth}})$	Expression	Numerical error
Mercury 0.330104 -2.8950	$-\left(e+\frac{e}{16}\right)$	0.0068
Venus 4.86732 -0.2046	$-\frac{e}{16}$	0.0347
Earth 5.97219 0.0000	$0e$	0.0000
Mars 0.641693 -2.2312	$-\left(\frac{e}{2}+\frac{e}{4}+\frac{e}{16}\right)$	0.0226
Ceres 0.000943 -8.7403	$-\left(2e+e+\frac{e}{8}+\frac{e}{16}\right)$	0.0758
Jupiter 1898.13 5.7615	$2e+\frac{e}{8}$	0.0148
Saturn 568.319 4.5556	$e+\frac{e}{2}+\frac{e}{8}+\frac{e}{16}$	0.0315
Uranus 86.8103 2.6766	e	0.0416
Neptune 102.410 2.8419	$e+\frac{e}{16}$	0.0463
Pluto 0.01309 -6.1193	$-\left(2e+\frac{e}{4}\right)$	0.0032
Sun 1989100 12.7161	$4e+\frac{e}{2}+\frac{e}{8}+\frac{e}{16}$	0.0258

3.2 Continued fraction analysis

Due to the fact that all the solar system data can be expressed by multiples of $\frac{e}{16}$, it is consistent to set all partial numerators in Müller's continued fractions (given in equation(2)) to Euler's number. We further follow the formalism of previous publications [5, 6] and introduce a phase shift p in equation (2). According to [6] the phase shift can only have the values 0 or ±1.5. So we write for instance for the masses of the celestial bodies:

$$\ln\frac{mass}{proton\ mass}=p+S \tag{3}$$

where S is the continued fraction

$$S=n_0+\cfrac{e}{n_1+\cfrac{e}{n_2+\cfrac{e}{n_3+...}}}. \tag{4}$$

We abbreviate $p+S$ as $[p; n_0 \mid n_1, n_2, n_3, \ldots]$. The free link n_0 and the partial denominators n_i are integers divisible by 3. For convergence reason, we have to include $|e{+}1|$ as allowed partial denominator. This means the free link n_0 is allowed to be $0, \pm 3, \pm 6, \pm 9 \ldots$ and all partial denominators n_i can take the values $e{+}1, -e{-}1, \pm 6, \pm 9, \pm 12 \ldots$.

Analogously we write for the planetary mean distances from the Sun:

$$\ln\frac{mean\ distance}{\lambda_C}=p+S \tag{5}$$

where $\lambda_C = \frac{h}{2\pi mc}$ is the reduced Compton wavelength of the proton with the numerical value $2.103089086\times 10^{-16}$ m. Since the exact diameter or radius of the proton is unknown, some other proton related parameter is used, which can be determined accurately. The same applies for the equatorial radii. For orbital and rotational periods we write:

$$\ln\frac{time\ period}{\tau}=p+S \tag{6}$$

where $\tau = \frac{\lambda_C}{c}$ is the oscillation period of a hypothetical photon with the reduced Compton wavelength of the proton and traveling with light speed (numerical value $7.015150081\times 10^{-25}$ s).

For the calculation of the continued fractions we did not consider any standard deviation of the published data. Practically, we developed the continued fraction and determined only 18 partial denominators. Next we calculated repeatedly the data value from the continued fraction, every time considering one more partial denominator. As soon as considering further denominators did not improve the experimental data value significantly (on the linear scale), we stopped considering further denominators and gave the resulting fraction in Tables 6-10. This means we demonstrate how accurately the published solar system data can be expressed through continued fractions. Additionally we gave also the numerical error, which is defined as absolute value of the difference between NASA's published data value and the value calculated from the continued fraction representation.

The continued fraction representations of the masses of celestial bodies are given in Table 6. As can be seen, the absolute value of the first partial denominator is frequently high, which locates the mass very close to the principal node.

Table 6: Continued fraction representation of masses (m) of celestial bodies according to equation (3) and absolute values of corresponding numerical errors.

Object m [kg]	Continued fraction representation Numerical error [kg]
Mercury 0.330104×10^{24}	[1.5; 114 | 9, -12, -e-1, e+1] $5.5e + 19$
Venus 4.86732×10^{24}	[1.5; 117 | -305223] 1.6×10^{14}
Earth 5.97219×10^{24}	[1.5; 117 | 12, e+1, -e-1, e+1] 3.0×10^{22}
Mars 0.641693×10^{24}	[0; 117 | -6, e+1, -6, 33, -60, -e-1, e+1, -e-1] 1.1×10^{15}
Ceres 9.43×10^{20}	[1.5; 108 | 6, 99, e+1, -e-1, e+1, -6, e+1, e-1] 3.4×10^{12}
Jupiter 1.89813×10^{27}	[1.5; 123 | -81, e+1, -e-1, -e-1, -e-1, e+1, -9, -e-1] 3.6×10^{18}
Saturn 5.68319×10^{26}	[0; 123 | 9, e+1, -e-1] 8.1×10^{24}
Uranus 8.68103×10^{25}	[1.5; 120 | -24, e+1, -e-1, e+1] 7.0×10^{22}
Neptune 1.0241×10^{26}	[1.5; 120 | 60, -e-1, e+1, -e-1] 3.9×10^{22}
Pluto 1.309×10^{22}	[1.5; 111 | 33, 9, -e-1, e+1, -18, e+1, e+1, -15] 3.2×10^{12}
Sun 1.9891×10^{30}	[0; 132 | -e-1, -e-1, e+1, -e-1, 12, -e-1] 5.0×10^{25} [1.5; 129 | e+1, -e-1, 15, e+1] 6.2×10^{26}

Table 7: Continued fraction representation of mean distances of celestial bodies from the Sun according to equation (5) and absolute values of corresponding numerical errors.

Object mean distance [km]	Continued fraction representation Numerical error
Mercury 57.91×10^6	[0; 60 | e+1, -e-1, -e-1, -e-1, 6, 6, -9, -e-1] 1 km
Venus 108.21×10^6	[1.5; 60 | 513, 6, -9, e+1] 260 m
Earth 149.595×10^6	[1.5; 60 | 9, -e-1, 51, e+1, 6, 6] 873 m
Mars 227.925×10^6	[0; 63 | -e-1, 30, -e-1, -15, 6, 9, -9] 0.4 m
Ceres 413.833×10^6	[0; 63 | -18, 9, e+1, -e-1, e+1, -e-1, e+1, -e-1] 5854 km
Jupiter 778.57×10^6	[0; 63 | 6, -9, 6, -e-1, e+1, -e-1, -6, 54] 372 m
Saturn 1433.525×10^6	[1.5; 63 | -6, -e-1, -e-1, -15, -48, e+1, -e-1] 8.7 km
Uranus 2872.46×10^6	no continued fraction found
Neptune 4495.06×10^6	[0; 66 | -e-1, 15, 15, 54, 9, -e-1, e+1, -e-1] 46 m [1.5; 63 | e+1, -597, -9, e+1] 181 km
Pluto 5906.375×10^6	[0; 66 | -6, 6, -e-1, -6, -15, -e-1, -12, -e-1] 7.2 km

In case of the Venus, the mass is almost exactly located in a node. Notably two low-weight bodies, Ceres and Mars, are most distant from the principal nodes. A preferred accumulation of planetary masses in nodes in agreement with results previously published by Müller [2]. This author published already a continued fraction analysis of planetary masses, however, the continued fractions were in the canonical form with all nominators equal 1. Interestingly, his result is principally not changed substituting the nominators for e. The only exception is the Sun, here even two continued fractions can be given and the mass is located in a non-turbulent zone between the principal nodes 129+1.5 and 132. This indicates that the probability of mass changes of the Sun is extremely low, so one can expect that all astrophysical parameters of the Sun will not show any evolution for a long time. We conclude that it seems to be a general property of mass to accumulate close to the nodes. Apparently no specific properties of the celestial bodies can be correlated to these data.

Table 7 displays the continued fraction representations of the mean distances from the Sun of the considered celestial bodies. When analyzing the denominators, it is directly clear that there is no general behavior of the planetary distances. For instance Venus is located almost in a node (n_1 very high), while Mercury, Mars and Neptune are far away from a node ($n_1 = e+1$ or $-e-1$). Uranus is even in a gap. Earth, Jupiter, Saturn and Pluto are moderately close to a node. This opens a door to associate a specific property of these bodies to the continued fraction representation. In this particular case we relate the mean distance to seismic activity of a solid object or heat release of a gas planet. The oscillation process inside Venus is turbulent, and it is known that Venus has an extreme

Table 8: Continued fraction representation of equatorial radii of celestial bodies according to equation (5) and absolute values of corresponding numerical errors.

Object Equatorial radius [km]	Continued fraction representation Numerical error
Mercury 2439.7	[0; 51 | -15, e+1, -e-1, e+1, -e-1] 1.6 km
Venus 6051.8	[0; 51 | e+1, 30, 9] 98 m
Earth 6378.14	[0; 51 | e+1, -15, -e-1, e+1, e+1] 57 m [1.5; 51 | -e-1, 207] 58 m
Mars 3397	[0; 51 | 21, -e-1, e+1, -e-1, e+1] 1.8 km
Ceres 487.3	[1.5; 48 | -9, 27, 9, 18] 0.01 m
Jupiter 71492	[0; 54 | 15, -18, -24, -6] 2 m
Saturn 60268	[0; 54 | 222, -6, -e-1] 46 m
Uranus 25559	[0; 54 | -e-1, 6, -e-1, -e-1, 9] 898 m [1.5; 51 | e+1, 6, 12, -e-1, e+1, -6] 44 m
Neptune 24764	[0; 54 | -e-1, e+1, e+1, 9, -e-1, 9] 22 m [1.5; 51 | e+1, e+1, 6, -6, -213] 0.05 m
Pluto 1151	[0; 51 | -e-1, e+1, -6, e+1, e+1] 475 m
Sun 6.955×10^5	[0; 57 | -6, e+1, -e-1, e+1, -e-1, -e-1, 12, -6] 49 m [1.5; 54 | e+1, -e-1, e+1, e+1, e+1 -e-1, e+1, -e-1, e+1] 21 km

volcanic activity [7,8]. Scientists also believe that the volcanism on Venus has been changing over time [7], so changes in trend may occur. The data also suggest that seismic activity on Earth is higher than on Mars, Mercury or Pluto.

For the gas planets Jupiter, Saturn and Neptune, it has been known that they produce more heat internally than they receive from the Sun [9, 10]. Contrary to this, Uranus is a relatively cold planet, radiating very little more energy than received. The principal source of this heating is believed to be a liberation of thermal energy from precipitation of Helium or other compounds in the interior of the planet while simultaneously gravitational potential energy is released.

Physically, such processes should exist in all gas planets, this means only the process kinetics can be associated to the continued fraction representation. We assume that the rate of this process is influenced by oscillations in the planet. For Uranus, which is located in a gap, the oscillation capability is low, which means the heat-releasing process occurred faster and is already almost completed. Jupiter and Saturn, located in proximity to the nodes 63 and 1.5+63, are in a fluctuation zone. So here the heat releasing process is disturbed and they are yet in a more early phase of process development, whereas Neptune (away from nodes) is in an already more advanced phase. From this we can predict that one day in future, first Neptune stops releasing excess heat, while Jupiter and Saturn will do this much later.

A very special situation is the continued fraction representation of dwarf planet Ceres. As can be seen, it has an exceptional high numerical error, actually this must be interpreted as "no continued fraction found". We report the fraction here only in order to demonstrate that the whole Asteroid belt is in a fluctuation zone around the node 63, which translates to $\lambda_C exp(63) = 3.22$ AU. This value is not acceptable as an average for the distances of the Asteroid belt objects from the Sun. Actually most Asteroids can be found between 2.1 and 3.5 AU. From this it can be concluded that most Asteroids accumulate in the compression zone before the principal node 63. Similarly is the situation for the Kuiper belt. All Kuiper belt objects are located before the node 66, $\lambda_C exp(66) = 64.77$ AU. The Astrophysics textbooks always teach the belt is located from the orbit of Neptune (30 AU) to 50 or 55 AU distance from the Sun. So again, the celestial bodies accumulate before a principal node.

Since Ceres is the largest Asteroid belt object, it is reasonable to claim Ceres is located in a gap, even inside a fluctuation zone. We interpret these fluctuations as the cause of the observed mass scattering in the whole Asteroid belt.

More research must still be done regarding the distribution of Kuiper belt objects. Brunini and Melita [11] suggested a Mars like object around 60 AU distance from the Sun in order to explain the Kuiper cliff, a sudden drop off of space rocks beyond 50 AU. Later, numerical simulations of Lykawka and Mukai showed that such a body would not reproduce the observed orbital distribution in the Kuiper belt [12], however these authors did not completely exclude the possibility of an unknown planet. Now, from our continued fraction analysis we suggest that there is indeed no unknown planet, it is just so that the compression zone before the principal node acts as accumulation site of these relatively light Kuiper belt objects. If there was such a solid planet in the fluctuation zone, it should possess volcanic activity similarly to Venus, and consequently should be very easy to detect, because of emission of infrared radiation. So this argument again confirms the absence of such a planet. Anyway, a detailed continued fraction analysis of Trans-Neptunian objects

combined with Kuiper belt objects would be very useful.

Table 8 displays analogously the continued fraction representations of planetary equatorial radii. From these data, some statements regarding the atmosphere of solid planets can be derived. We interprete an atmosphere as an extension of a planet with the effect to increase its radius. On the other hand, an atmosphere is also governed by the chemical composition of a planet and its temperature and these parameters are more decisive. Such an analysis cannot be applied to gaseous planets, since they always have a very dense atmosphere, regardless of their radii.

The most dense atmospheres can be found on Earth and on Venus. The first partial denominator in the continued fraction representation of Venus is e+1. this means the radius of Venus is in an expansion zone and far away from the node. An increase in radius is favored and any probabilities of trend changes are low. This is in agreement with the observed high density of the atmosphere on Venus, with a pressure of 95 bar at the surface [8]. In the case of our planet Earth, two continued fractions can be given, so the radius is influenced by the two nodes 51 and 51+1.5. Both first partial denominators put the radius far away from the corresponding nodes into a non-fluctuation zone. Here does not exist any specific trend and the formation of the atmosphere is solely governed by chemical composition and temperature.

Pluto is with a negative first partial denominator in a compression zone, so the expansion of its radius by an atmosphere is not favored. Indeed Pluto has only a very thin atmosphere in the micro-bar range [13]. According to reference [14], Pluto's atmosphere at perihelion extends to depths greater than Earth's atmosphere and may even enclose the moon Charon. The atmosphere is thought to be actively escaping, so Pluto is the only planet in the solar system actively losing its atmosphere now.

The same is true for Mercury. In agreement with the observations, Mercury does not have an atmosphere [8], which can also be alternatively explained by its high surface temperature.

Mars is with the positive number 21 of the first partial denominator in an expansion zone, so the formation of an atmosphere is favored. At the same time the radius is also close to the node 51 in a fluctuation zone. This means changes in process trends may occur. Considering the formation of an atmosphere as the relevant process, this process can be interrupted or inverted over long time periods. As a consequence, one would expect an atmosphere, but significantly thinner than that on Venus. Actually the surface pressure on Mars is close to 1% to that of the Earth and there are speculations that the atmosphere on Mars has experienced major changes in the past [8].

Ceres is a low density object consisting of rock and ice with mean density of only 2 g/cm^3, which supports the presence of a lot of ice. The "frost line" in our solar system — the distance where ice will not evaporate — is roughly at 5 AU from the Sun [15]. So one must ask why Ceres does not have already lost all his ice through sublimation. From the continued fraction representation, the radius of Ceres is in a compression zone and the formation of an atmosphere is not favored. Through evaporation of the ice, at least temporarily an atmosphere will form. For this reason we believe Ceres is able to continue for a long time as an icy dwarf planet.

When looking at the data it turns out that the gaseous planets seem to prefer radii that can be described by two con-

Table 9: Continued fraction representation of sidereal orbital periods of celestial bodies according to equation (6) and absolute values of corresponding numerical errors.

Planet T [s]	Continued fraction representation Numerical error
Mercury 7595370	[0; 72 | -6, e+1, -e-1, e+1, -30, -e-1, -33, -6] 0.002 s [1.5; 69 | e+1, -e-1, e+1, 6, -12, 6, -e-1, e+1, -15] 0.01 s
Venus 19400861	[0; 72 | 6, e+1, -6, 6, e+1, -e-1, e+1, -e-1, e+1] 128 s
Earth 31536549	[0; 72 | e+1, -e-1, -6, e+1, -6, -6, -e-1, 9, -6] 0.1 s [1.5; 72 | -e-1, -e-1, -12, 45, e+1, -6, -e-1, -24] 0.0003 s
Mars 59314410	[1.5; 72 | 183, -e-1, 12, -e-1, e+1, -e-1] 13 s
Ceres 145065600	[0; 75 | -e-1, -e-1, e+1, 6, -e-1, 6, -6, -18, e+1] 0.3 s [1.5; 72 | e+1, -e-1, -225, -e-1, e+1, -e-1, -9, -e-1] 0.06 s
Jupiter 374099427	no continued fraction found
Saturn 928656297	[1.5; 75 | -12, 6, e+1, -e-1, 33, e+1, -e-1, e+1, -e-1] 74 s
Uranus 2649555255	[0; 78 | -e-1, -12, e+1, -e-1, 12, -e-1, -69, -9] 0.9 s
Neptune 5196859068	[0; 78 | -225, e+1, -9, e+1, -6, e+1, 48] 0.04 s
Pluto 7818425618	no continued fraction found

tinued fractions. For the Sun, Uranus and Neptune it can be said that they are influenced by two neighbored nodes. This indicates their sizes will remain constant over a longe time. The only exceptions are Jupiter and Saturn, which are in an expansion zone. One would expect their sizes increasing. How could this be achieved in practice? There is only one possibility, Jupiter and Saturn must capture some asteroids or comets preferentially from the Kuiper belt. When looking at the number of their moons, it can be assumed that such a process has already been progressing for a long time. A moon can be interpreted as an incomplete capture, this means the object was captured without crashing into the planet and increasing its size. Indeed Jupiter and Saturn have 63 and 62 confirmed moons, while Uranus has 27, and Neptune only 13 moons. Normally one would expect that Uranus and Neptune should have the most moons, since they are much closer located to the Kuiper belt. Notably 55 of Jupiter's moons are irregular satellites with high eccentricities and inclinations, while Saturn has just 38 of such satellites. It is assumed that these irregular satellites were captured from other orbits.

In Table 9, the continued fraction representations of the orbital periods are given. When analyzing these fractions, their interpretation is problematic: One has to bear in mind that Kepler's 3rd law relates the orbital period to the semi-major axis (for most planets close to the mean distance), so these parameters are not independent from each other.

Regarding oscillation properties, it is clearly visible that the continued fraction representations of the orbital periods do not provide a similar image of planetary features than the representations of the corresponding mean distances. For instance, the orbital periods of Mars and Neptune are located in a highly turbulent zone. This is contrary to to the continued fraction representation of its mean distances given in Table 7, where both planets are far away from a node. Since for the mean distances a meaningful continued fraction representation exists, the orbital periods do not fit anymore in this model and their mathematical representation in continued fractions, as presented here, is physically meaningless.

Luckily, the situation is easier for the rotation periods of the celestial bodies (see Table 10). As can be seen, the rotation periods prefer values far away from the nodes in non-fluctuating zones. There are only three exceptions: Jupiter Saturn and Ceres have periods located in a principal node. This means the rotation periods are in an early stage of development, which can be justified with a specific process inside the celestial bodies.

For the gas planets Jupiter and Saturn it has been known that heat is generated from precipitation of Helium or other compounds in the interior of the planet while simultaneously gravitational potential energy is released. Through such a process, the moment of inertia of the planet changes gradually and the rotation period evolves. From the analysis of the mean distances of Jupiter and Saturn, we have already stated that their heat release processes are still in an early phase of development. Exactly the same can be derived from the analysis of rotation periods. The rotation of the Sun is also not yet completely evolved, however here this effect is minor. Any internal structuring of plasma fluxes could be responsible for this.

Table 10: Continued fraction representation of sidereal rotation periods (T) of celestial bodies according to equation (6) and absolute values of corresponding numerical errors.

Planet T [s]	Continued fraction representation Numerical error
Mercury 5067032	[0; 72 | -e-1, e+1, -6, 6, -15, -e-1, e+1, -e-1, e+1] 3 s
Venus 20996755	[0; 72 | 6, -9, -12, 18, -9, e+1] 0.1 s
Earth 86164	[0;66 | e+1, -e-1, e+1, -6, e+1, e+1, -e-1, 21] 0.02 s [1.5; 66 | -6, e+1, -15, -e-1, -6] 0.07 s
Mars 88643	[1.5; 66 | -6, 6, -18, -12] 0.04 s
Ceres 32668	[0; 66 | 255, -e-1, e+1, -e-1, e+1] 0.17 s
Jupiter 35730	[0; 66 | 27, 27, -21] 0.005 s
Saturn 38362	[0; 66 | 15, e+1, -e-1, -e-1, e+1, -e-1] 2 s
Uranus 62064	[0; 66 | e+1, 6, 39, -12] 0.02 s [1.5; 66 | -e-1, 6, -e-1, -9, -e-1, 48] 0.001 s
Neptune 57996	[0; 66 | e+1, e+1, -e-1, 9, -9, -18] 0.003 s [1.5; 66 | -e-1, e+1, -30, -e-1, e+1, -e-1] 4 s
Pluto 551854	no continued fraction found
Sun 2164320	[1.5; 69 | -9, -15, e+1, e+1, -6, 9] 0.003 s

Ceres has an unusual location inside the Asteroid belt, which is a turbulent zone as can be derived from the continued fraction analysis of its mean distance from the Sun. Knowing this, we speculate that the evolution of its rotation period could have been influenced by the fluctuating population of the belt through collisions of an early Ceres with many smaller asteroids over a long time. According to reference [15], there are possibly volatile compounds in the interior of Ceres. Ceres could have accreted from rocky and icy planetesimals. This has taken some time, we speculate that

possibly Ceres had less time for the evolution of its rotation than other planets.

An other reference [16] speculates regarding a subsurface ocean and mentions a modeling predicting that ice in the outer 10 km of Ceres would always remain frozen, although the frozen crust would be gravitationally unstable and likely overturn, melt, and re-freeze. Such repeatedly occurring movements of heavy masses on Ceres could have interfered with the evolution of its rotation period.

4 Conclusions

Numerical investigation of solar system data revealed that masses, radii, distances of celestial bodies from the Sun, orbital periods and rotation periods can be expressed as multiples of $\frac{e}{16}$ on the logarithmic number line, which proves that they are not a set of random numbers. Through application of a fractal scaling model, we set these numerical values in relation to proton resonances and correlated numerous features of celestial bodies with their oscillation properties. From this it can be concluded that the continued fraction representations with all nominators equal e are adequate and Müller's fractal model turned out to be a powerful tool to explain the fractal nature of the solar system. If some day in future, a further planet will be discovered in our solar system, it should be possible to derive analogously some of its features from its orbital parameters.

Acknowledgments

The authors greatly acknowledge the financial support from the Brazilian governmental funding agencies FAPESQ and CNPq.

Submitted on December 26, 2010 / Accepted on December 27, 2010

References

1. Müller H. Fractal scaling Models of resonant oscillations in chain systems of harmonic oscillators. *Progress in Physics*, 2009, v. 2 72–76.
2. Müller H. Fractal scaling models of natural oscillations in chain systems and the mass distribution of the celestial bodies in the solar system. *Progress in Physics*, 2010, v. 1 62–66.
3. Müller H. Fractal scaling models of natural oscillations in chain systems and the mass distribution of particles. *Progress in Physics*, 2010, v. 3 61–66.
4. Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P. Numerical recipes in C. Cambridge University Press, Cambridge, 1992.
5. Ries A., Fook M. V. L. Fractal structure of nature's preferred masses: Application of the model of oscillations in a chain system. *Progress in Physics*, 2010, v. 4, 82–89.
6. Otte R., Müller H.. German patent No. DE102004003753A1, date: 11.08.2005
7. Basilevsky A. T., Head J. W. Venus: Timing and rates of geologic activity. *Geology*, 2002, v. 30, no. 11, 1015-–1018.
8. Gregersen E. (Editor). The inner solar system: the Sun, Mercury, Venus, Earth, and Mars. Britannica Educational Publishing, New York, 2010.
9. Elkins-Tanton L. T. Jupiter and Saturn. Chelsea House, New York, 2006.
10. Gregersen E. (Editor). The outer solar system: Jupiter, Saturn, Uranus, Neptune, and the dwarf planets. Britannica Educational Publishing, New York, 2010.
11. Brunini A., Melita M. D. The Existence of a Planet beyond 50 AU and the Orbital Distribution of the Classical Edgeworth-Kuiper-Belt Objects. *Icarus*, 2002, v. 160, no. 1, 32–43.
12. Patryk S. Lykawka P. S., Mukai T. An Outer Planet beyond Pluto and the Origin of the Trans-Neptunian belt Architecture. *The Astronomical Journal*, 2008, v. 135, 1161-–1200.
13. Lellouch E., Sicardy B., de Bergh C., Käufl H.-U., Kassi S., Campargue A. Pluto's lower atmosphere structure and methane abundance from high-resolution spectroscopy and stellar occultations. arXiv:0901.4882v1, [astro-ph.EP], 30 Jan 2009.
14. Elkins-Tanton L. T. Uranus, Neptune, Pluto, and the Outer Solar System. Chelsea House, New York, 2006.
15. Carry B., Dumas C., Fulchignoni M., Merline W. J., Berthier J., Hestroffer D., Fusco T., Tamblyn P. Near-Infrared Mapping and Physical Properties of the Dwarf-Planet Ceres. arXiv:0711.1152v1, [astro-ph], 7 Nov 2007.
16. Rivkin A. S., Volquardsen E. L., Clark B. E. The surface composition of Ceres: Discovery of carbonates and iron-rich clays. *Icarus*, 2006, v. 185, 563–567.

Photon-Assisted Resonant Chiral Tunneling Through a Bilayer Graphene Barrier

Aziz N. Mina* and Adel H. Phillips†
*Faculty of Science, Beni-Suef University, Beni-Suef, Egypt
†Faculty of Engineering, Ain-Shams University, Cairo, Egypt

E-mail: adel_phillips@yahoo.com

The electronic transport property of a bilayer graphene is investigated under the effect of an electromagnetic field. We deduce an expression for the conductance by solving the Dirac equation. This conductance depends on the barrier height for graphene and the energy of the induced photons. A resonance oscillatory behavior of the conductance is observed. These oscillations are strongly depends on the barrier height for chiral tunneling through graphene. This oscillatory behavior might be due to the interference of different central band and sidebands of graphene states. The present investigation is very important for the application of bilayer graphene in photodetector devices, for example, far-infrared photodevices and ultrafast lasers.

1 Introduction

Two-dimensional graphene monolayer and bilayer exhibit fascinating electronic [1–4] and optical properties [5, 6] due to zero energy gap and relativistic-like nature of quasiparticle dispersion close to the Fermi-level. With recent improvements in nanofabrication techniques [7] the zero-energy gap of graphene can be opened via engineering size, shape, character of the edge state and carrier density, and this in turn offers possibilities to simultaneously control electronic [8, 9] magnetic [10, 11] and optical [6, 12] properties of a single material nanostructure. Recent studies have also addressed electronic properties of confined graphene structure like dots, rings or nanoribbons. In particular, nanoribbons have been suggested as potential candidates for replacing electronic components in future nanoelectronic and spintronic devices [3, 13]. Recent research shows that graphene [14] is a suitable candidate to examine the photon-assisted tunneling and quantum pumps in the Dirac system.

The purpose of the present paper is to investigate the angular dependence of the chiral tunneling through double layer graphene under the effect of the electromagnetic field of wide range of frequencies.

2 Theoretical Formulation

In this section, we shall derive an expression for the conductance of a bilayer graphene by solving the eigenvalue problem Dirac equation. The chiral fermion Hamiltonian operates in space of the two-component eigenfunction, ψ, where Dirac eigenvalue differential equation is given by [14, 15]:

$$-i v_F \vec{\sigma} \cdot \vec{\nabla} \psi(r) = E\psi(r)\,, \tag{1}$$

where $\vec{\sigma}$ are the Pauli-matrices, V_F is the Fermi-velocity, and E is the scattered energy of electrons. It is well known that graphene junction have finite dimensions [14, 15], the motion of chiral fermions is quantized. This quantization imposes additional constrains on the directional tunneling diagram. So, accordingly, the value of the angle of incidence of electrons on the barrier could be obtained from boundary conditions along the y-direction as we will see below.

In order to solve Eq.(1), we propose a potential barrier of width, L, and height, V_0,. The eigenfunction, $\psi_L(r)$ in the left of the potential barrier is given by:

$$\psi_L(r) = \sum_{n=-\infty}^{\infty} J_n\left(\frac{eV_{ac}}{\hbar\omega}\right)\left\{e^{[i(k_x x+k_y y)]} + \right.$$
$$\left. + \frac{R_n(E)}{\sqrt{2}}\begin{pmatrix}1\\ s\, e^{i(\pi-\phi)}\end{pmatrix} e^{[i(-ik_x x+k_y y)]}\right\}, \tag{2}$$

where the angle $\phi = \tan^{-1}(\frac{k_y}{k_x})$, in which $k_x = k_f\cos(\phi)$ and $k_v = k_F\sin(\phi)$,and k_F is the Fermi-wave number, and J_n is the n^{th} order Bessel function, V_{ac} is the amplitude of the induced photons of the electromagnetic field with frequency, ω, and $R_n(E)$ is the energy-dependent reflection coefficient.

The eigenfunction, $\psi_b(r)$, inside the potential barrier is given by:

$$\psi_b(r) = \sum_{n=-\infty}^{\infty} J_n\left(\frac{eV_{ac}}{\hbar\omega}\right)\left\{\frac{a}{\sqrt{2}}\begin{pmatrix}1\\ s'\, e^{i\theta}\end{pmatrix} e^{[i(q_x x+k_y y)]} + \right.$$
$$\left. + \frac{b}{\sqrt{2}}\begin{pmatrix}1\\ s'\, e^{i(\pi-\theta)}\end{pmatrix} e^{[i(-q_x x+k_y y)]}\right\}, \tag{3}$$

where the angle $\theta = \tan^{-1}(\frac{k_y}{q_x})$, and the wave number qx is expressed as:

$$q_x = \sqrt{\frac{(V_0-\varepsilon)^2}{v_F^2} - k_y^2} \tag{4}$$

and $\varepsilon = E - eV_g - \hbar\omega$, V_0 is the barrier height, E is the energy of the scattered electrons, V_g is the gate voltage and $\hbar\omega$ is the photon energy.

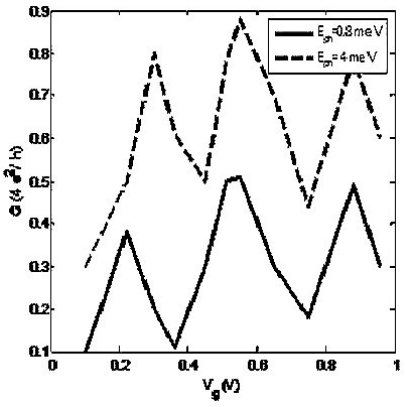


Fig. 1: The variation of the conductance, G, with gate voltage V_g, at different photon energies, E_{ph}.

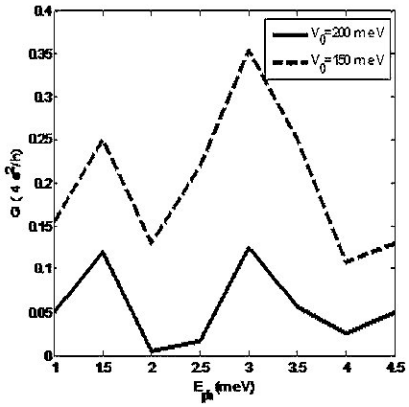


Fig. 2: The variation of the conductance, G, with the photon energy, E_{ph}, at different values of barrier height, V_0.

The eigenfunction, $\psi_R(r)$, in the right region to the potential barrier which represents the transmitted electrons is given by:

$$\psi_R(r) = \sum_{n=-\infty}^{\infty} J_n\left(\frac{eV_{ac}}{\hbar\omega}\right)\left\{\frac{\Gamma_n(E)}{\sqrt{2}}\begin{pmatrix}1\\ s\,e^{i\theta}\end{pmatrix}e^{[i(k_x x+k_y y)]}\right\}, \quad (5)$$

where $\Gamma_n(E)$ are the transmitted electron waves through the barrier. The parameters s and s' are expressed as:

$$s = \mathrm{sgn}\,(E) \qquad \text{and} \qquad s' = \mathrm{sgn}\,(E - V_0)\,. \quad (6)$$

Now, the coefficients $R_n(E), a, b, \Gamma_n(E)$ could be determined by applying the continuity conditions of the eigenfunctions, Eqs.(2,3,5), at the boundaries as follows:

$$\left.\begin{aligned}&\psi_L\,(x=0,\,y) = \psi_b\,(x=0,\,y)\\ &\text{and}\\ &\psi_b\,(x=L,\,y) = \psi_R\,(x=L,\,y)\end{aligned}\right\}. \quad (7)$$

So, the transmission probability, $|\Gamma_n(E)|^2$, could be determined from the boundary conditions Eq.(7) and is given by:

$$|\Gamma_n(E)|^2 = \sum_{n=-\infty}^{\infty} J_n^2\left(\frac{eV_{ac}}{\hbar\omega}\right)\times$$
$$\times\left\{\frac{\cos^2(\theta)\cos^2(\phi)}{[[\cos(Lq_x)\cos\phi\cos\theta]^2+\sin^2(Lq_x)(1-ss'\sin\phi\sin\theta)^2]}\right\}. \quad (8)$$

The conductance, G, is given by [16, 17]:

$$G(E) = \frac{4e^2}{h}\int dE\,|\Gamma_n(E)|^2\left(-\frac{\partial f_{FD}}{\partial E}\right), \quad (9)$$

where f_{FD} is the Fermi-Dirac distribution function. Now, substituting Eq.(8) into Eq.(9), we get a complete expression for conductance which depends on the angles ϕ, θ, and on the barrier height, V_0, and its width, the gate voltage, V_g, and the photon energy, $\hbar\omega$.

3 Results and Discussions

The conductance, G, has been computed numerically as a function of the gate voltage, V_g, and photon energy, $E_{ph} = \hbar\omega$ of the induced electromagnetic field. For the bilayer graphene, the effective mass of the fermion quasiparticle m^* equals approximately 0.054 m_e [14, 15]. The parameter m_e is the free mass of the electron. The main features of the present results are:

(1) Fig.(1) shows the variation of the conductance, with the gate voltage, V_g, at different values of the photon energies of the induced electromagnetic field. We notice an oscillatory behavior of the conductance. The electromagnetic field induces resonant peaks in the photon-assisted chiral tunneling conductance.

(2) Fig.(2) shows the dependence of the conductance on the energy of the induced photons at different values of the barrier height, V_0. An oscillation of the conductance is observed.

The observed oscillations in conductance for Figs.(1,2) can be explained as Follows: For grapheme under the effect of the electromagnetic field, the chiral tunneling of electrons can undergo transitions between the central band to several sidebands by means of photon emission or absorption. Such process is referred to as photo-assisted tunneling [18–20]. Also, the phase correlations during chiral tunneling can be directly tuned by applying of an external electromagnetic field leads to a resonance trend in the conductance of a bilayer graphene.

The present results show a good concordant with those in the literature [21–23].

4 Conclusion

The present investigation shows that the chiral tunneling of Dirac electrons through graphene enables ultra-wide band tunability. The rise of graphene in photonics and optoelectronics is shown by several results ranging from photo-detectors, light emitted devices, solar cells and ultra-fast lasers [23, 24].

Submitted on December 12, 2010 / Accepted on December 19, 2010

References

1. Novoselov K.S., Geim A.K., Morozov S.V., Jiang D., Katsnelson M.I., Grigoriev I.V., Dubonos S.V. and Firsov A.A. Two-dimensional gas of massless Dirac fermions in graphene. *Nature*, 2005, v.438, 197.
2. Zhang Y.B., Tan Y.W., Stormer H.L., and Kim P. Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature*, 2005, v.438, 2001.
3. Geim A.K. and Novoselov K.S. The rise of graphene. *Nature Materials*, 2007, v.6, 183.
4. Tworzydlo J., Trauzettel B., Titov M. Rycerz A. and Beenakker C.W.J. Quantum limited shot noise in graphene. *Physical Review Letters*, 2006, v.96, 246802.
5. Sadowski M.L., Martinez G., Potemski M., Berger C. and de Haar W.A. Landau level spectroscopy of ultra thin graphite layers. *Physical Review Letters*, 2006, v.97, 266405.
6. Mueller T., Xia F. N. A., and Avouris P. Graphene photo-detectors for high speed optical communications. *Nature Photonics*, 2010, v.4, 297.
7. Campos L.C., Manfrinato V. R., Sanchez-Yamagishi J.D., Kong J., and Jarillo-Herrero P. Anisotropic etching and nanoribbon formation in single layer graphene. *Nano Letters*, 2009, v.9, 2600.
8. Libish F., Stampfer C., and Bugdorfer J. Graphene quantum dots: beyond a Dirac billiard. *Physical Review B*. 2009, v.79, 115423.
9. Ezawa M. Dirac fermions in a graphene nanodisk and a graphene corner: Texture of vortices with an unusual winding number. *Physical Review B*, 2010, v.81, 201402.
10. Ezawa M. Metallic graphene nanodisks: Electrical and magnetic properties. *Physical Review B*, 2007, v.76, 245415.
11. Wang W. L., Meng S. and Kaxiras E. Graphene nanoflakes with large spin. *Nano Letters*, 2008, v.8, 241.
12. Yan X., Cui X., Li B.S., and Li L.S. Large solution processable graphene quantum dots as light absorber for photovoltaics. *Nano Letters*, 2010, v.10, 1869.
13. Son J.W., Cohen M. L. and Louie S. G. Half-metallic graphene nanoribbons. *Nature*, 2006, v.444, 347.
14. Castro Neto A.H., Guinea F., Peres N.M.R., Novoselov K.S. and Geim A.K. The electronic properties of graphene. *Reviews of Modern Physics*, 2009, v.81, 109.
15. Beenakker C.W. J. Andreev reflection and Klein tunneling in graphene. *Reviews of Modern Physics*, 2008,v.80, 1337.
16. Datta S. Electronic transport in mesoscopic systems. Cambridge University Press, 1997, Cambridge.
17. McEuen P.L., Bockrath M., Cobden D.H., Yoon Y.G., and Louie S.G. Disorder, pseudospin, and back scattering in carbon nanotubes. *Physical Review Letters*, 1999, v.83, 5098.
18. Platero G. and Aguado R. Photon-assisted transport in semiconductor nanostructures. *Physics Reports*, 2004, v.394, 1.
19. Mina A.N., Awadalla A.A., Phillips A.H., and Ahmed R.R.. Microwave spectroscopy of carbon nanotube field effect transistor. *Progress in Physics*, 2010, v.4, 61.
20. Awadalla A.A., Phillips A.H., Mina A.N., and Ahmed R.R. Photon-assisted transport in carbon nanotube mesoscopic device. Accepted for publication in *International Journal of Nanoscience*, 2010.
21. Shafranjuk S. Probing the intrinsic state of a one-dimensional quantum well with photon-assisted tunneling. *Physical Review B*, 2008, v.78, 235115.
22. Abergel D.S.L. and Chakraborty T. Generation of valley polarized current in bilayer graphene. *Applied Physics Letters*, 2009, v.95, 062107.
23. Rocha C.G., Foa Terres L.E. F., and Cuniberti G. AC-transport in graphene-base Fabry-Perot devices. *Physical Review B*, 2010, v.81, 115435.
24. Bonaccorso F., Sun Z., Hasan T., and Ferrari A.C. Graphene photonics and optoelectronics. *Nature Photonics*, 2010, v.4, 611.

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), Zentralblatt MATH (Germany), 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,
200 College Road, 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 2011 **VOLUME 2**

CONTENTS

Information for Authors and Subscribers

Progress in Physics has been created for 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. 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. We accept brief papers, no larger than 8 typeset journal pages. Short articles are preferable. Large papers can be considered in exceptional cases to the section *Special Reports* intended for such publications in the journal. 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.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

The Lorentz Transformation as a Planck Vacuum Phenomenon in a Galilean Coordinate System

William C. Daywitt
National Institute for Standards and Technology (retired), Boulder, Colorado, USA
E-mail: wcdaywitt@earthlink.net

In a seminal Masters' dissertation [1] Pemper derived the relativistic electric and magnetic fields of a uniformly moving charge from the response of some continuum to the perturbation from the charge's Coulomb field. The results seem to imply that the Maxwell equations and the Lorentz transformation are associated with some type of vacuum state. Unbeknownst at the time, Pemper had discovered the Planck vacuum (PV) quasi-continuum [2] and its interaction with the free charge. The importance of this derivation, its obscurity in the literature, and its connection to the PV justifies the following rework of that derivation.

1 Pemper Derivation

When a free, massless, bare charge e_* travels in a straight line at a uniform velocity v its bare Coulomb field e_*/r^2 perturbs (polarizes) the PV [2]. If there were no PV, the bare field would propagate as a frozen pattern with the same velocity and there would be no accompanying magnetic field. The corresponding force perturbing the PV is e_*^2/r^2, where one of the charges e_* in the product e_*^2 belongs to the free charge and the other to the individual Planck particles making up the degenerate negative-energy PV.

This charge-vacuum interaction is described by Pemper [1] as a series ($n = 1, 2, 3, \ldots$) of electric and magnetic fields (*generated by the vacuum*)

$$\nabla \times \mathbf{E}_n = -\frac{1}{c}\frac{\partial \mathbf{B}_n}{\partial t} \tag{1}$$

and

$$\mathbf{B}_{n+1} = \boldsymbol{\beta} \times \mathbf{E}_n \tag{2}$$

that respond in a iterative fashion to the bare charge's Coulomb field, leading to the well-known relativistic electric and magnetic fields that are traditionally ascribed to the charge as a single entity. The serial electric and magnetic fields are $\mathbf{E}_n$ and $\mathbf{B}_n$ and $\boldsymbol{\beta} = \mathbf{v}/c$. The curl equation in (1) is recognized as the Faraday equation and the magnetic field in (2) is due to the free-charge field rotating the induced dipoles within the PV. The series of partial fields is not envisioned as a series in time — the PV response is assumed to happen instantaneously at each field point.

The initial magnetic field in the series is $\mathbf{B}_1 = \boldsymbol{\beta} \times \mathbf{E}_0$, where the bare charge's laboratory-observed Coulomb field is

$$\mathbf{E}_0 = \frac{e\mathbf{r}}{r^3} = \frac{e}{e_*}\frac{e_*\mathbf{r}}{r^3} = \alpha^{1/2}\frac{e_*\mathbf{r}}{r^3}, \tag{3}$$

where α is Planck's constant. The serial electric fields are assumed to be radial; so the final electric field is radial with a magnitude equal to the sum

$$E = E_0 + E_1 + E_2 + E_3 + \ldots, \tag{4}$$

where the E_n are the magnitudes of the $\mathbf{E}_n$s and the final magnetic field is $\boldsymbol{\beta} \times \mathbf{E}$. Assuming that the $E_n = E_n(r, \theta)$, the charge-PV feedback equations (1) and (2) reduce to

$$\frac{\partial E_n}{\partial \theta} = \frac{r}{c}\frac{\partial B_n}{\partial t} \tag{5}$$

and

$$B_{n+1} = \beta E_n \sin\theta \tag{6}$$

in the azimuthal direction about the z-axis.

Calculating the first partial field E_1 in the series begins with (6)

$$B_1 = \beta E_0 \sin\theta \tag{7}$$

and leads to (Appendix A)

$$\dot{B}_1 = \frac{3c\beta^2 E_0 \sin\theta\cos\theta}{r}, \tag{8}$$

where the overhead dot represents a partial differentiation with respect to time. Then from (5)

$$dE_1 = \frac{r\dot{B}_1}{c}\,d\theta = 3\beta^2 E_0 \sin\theta\cos\theta\,d\theta, \tag{9}$$

which integrates over the limits $(0, \theta)$ to

$$E_1 = \frac{3\beta^2 E_0 \sin^2\theta}{2} - \lambda_1 E_0, \tag{10}$$

where the reference field $E_1(\theta = 0) = -\lambda_1 E_0$ with λ_1 a constant to be determined.

The second iteration for the electric field begins with

$$B_2 = \beta E_1 \sin\theta = \frac{3\beta^3 E_0 \sin^3\theta}{2} - \lambda_1 B_1 \tag{11}$$

and yields (Appendix A)

$$\dot{B}_2 = \frac{15c\beta^4 E_0 \sin^3\theta\cos\theta}{2r} - \lambda_1 \dot{B}_1\,. \tag{12}$$

Equation (5) then leads to

$$dE_2 = \frac{r\dot{B}_2}{c}\,d\theta = \left(\frac{15\beta^4 E_0 \sin^3\theta\cos\theta}{2} - \frac{\lambda_1 r\dot{B}_1}{c}\right)d\theta, \quad (13)$$

which integrates to

$$E_2 = \frac{15\beta^4 E_0 \sin^4\theta}{8} - \lambda_1\frac{3\beta^2 E_0 \sin^2\theta}{2} - \lambda_2 E_0, \quad (14)$$

where again $E_2(\theta = 0) = -\lambda_2 E_0$.

The third iteration proceeds as before and results in (Appendix A)

$$\dot{B}_3 = \frac{3\cdot 5\cdot 7c\beta^6 E_0 \sin^5\theta\cos\theta}{8r} - \lambda_1\frac{3\cdot 5c\beta^4 E_0 \sin^3\theta\cos\theta}{2r}$$
$$-\lambda_2\frac{3c\beta^2 E_0 \sin\theta\cos\theta}{r} \quad (15)$$

and

$$E_3 = \frac{3\cdot 5\cdot 7\beta^6 E_0 \sin^6\theta}{6\cdot 8} - \lambda_1\frac{3\cdot 5\beta^4 E_0 \sin^4\theta}{2\cdot 4}$$
$$-\lambda_2\frac{3\beta E_0 \sin^2\theta}{2} - \lambda_3 E_0 \quad (16)$$

for the third partial field.

Inserting (10), (14), and (16) (plus the remaining infinity of partial fields) into (4) gives

$$E = E_0 + \frac{3\beta^2 E_0 \sin^2\theta}{2} + \frac{3\cdot 5\beta^4 E_0 \sin^4\theta}{8}$$
$$+\frac{3\cdot 5\cdot 7\beta^6 E_0 \sin^6\theta}{48} + \ldots$$
$$-\lambda_1\left(E_0 + \frac{3\beta^2 E_0 \sin^2\theta}{2} + \frac{3\cdot 5\beta^4 E_0 \sin^4\theta}{8} + \ldots\right)$$
$$-\lambda_2\left(E_0 + \frac{3\beta E_0 \sin^2\theta}{2} + \ldots\right) - \lambda_3(E_0 + \ldots) + \ldots$$
$$= E_0\left(1 + \frac{3\beta^2 \sin^2\theta}{2} + \frac{3\cdot 5\beta^4 \sin^4\theta}{2\cdot 4}\right.$$
$$\left.+\frac{3\cdot 5\cdot 7\beta^6 \sin^6\theta}{2\cdot 4\cdot 6} + \ldots\right)(1-\lambda), \quad (17)$$

where

$$\lambda \equiv \sum_{n=1}^{\infty}\lambda_n \quad (18)$$

is a constant. The sum after the final equal sign in (17) is recognized as the function $(1-\beta^2\sin^2\theta)^{-3/2}$; so E can be expressed as

$$E = \frac{(1-\lambda)E_0}{(1-\beta^2\sin^2\theta)^{3/2}}\,. \quad (19)$$

Finally, the constant λ can be evaluated from Gauss' law and the conservation of bare charge e_*:

$$\int \mathbf{D}\cdot\mathbf{dS} = 4\pi e_* \longrightarrow \int \mathbf{E}\cdot\mathbf{dS} = 4\pi e, \quad (20)$$

where $\mathbf{D} = (e_*/e)\mathbf{E}$ is used to arrive at the second integral. Inserting (19) into (20) and integrating yields

$$\lambda = \beta^2, \quad (21)$$

which, inserted back into (19), gives the relativistic electric field of a uniformly moving charge. That this field is the same as that derived from the Lorentz transformed Coulomb field is shown in Appendix B.

2 Conclusions and Comments

The calculations of the previous section suggest that the Lorentz transformation owes its existence to interactions between free-space particles and the negative-energy PV. Free space is defined here as "the classical void + the zero-point electromagnetic vacuum" [3].

The fact that the bare charge is massless makes the Pemper derivation significantly less involved and more straightforward than the related case for the massive point charge (Dirac electron). Nevertheless, the uniform motion of the Dirac electron too exhibits electron-PV effects. When a bare charge is injected into free space (presumably from the PV) it very quickly ($\sim 10^{-30}$ sec) develops a mass from being driven by the random fields of the electromagnetic vacuum. The corresponding electron-PV connection is easily recognized in the Lorentz-covariant Dirac equation [4, p. 90], [5]:

$$\left(ic\hbar\gamma^\mu\partial_\mu - mc^2\right)\psi = 0 \longrightarrow \left(ie_*^2\gamma^\mu\partial_\mu - mc^2\right)\psi = 0, \quad (22)$$

where the PV relation $c\hbar = e_*^2$ is used to arrive at the equation on the right. A nonrelativistic expression for the electron mass is given by Puthoff [3, 6]

$$m = \frac{2}{3}\frac{\left\langle \dot{r}^2\right\rangle^{1/2}}{c}\,m_*, \quad (23)$$

where $\dot{\mathbf{r}}$ represents the random excursions of the zero-point-driven bare charge about its center of (random) motion at $\mathbf{r} = 0$ and m_* is the Planck mass.

The massive point charge perturbs the PV with the two-fold force [5]

$$\frac{e_*^2}{r^2} - \frac{mc^2}{r}, \quad (24)$$

where the first and second terms are the polarization and curvature* forces respectively. It is the interaction of this composite force with the PV that is responsible for the Dirac equation as evidenced by the e_*^2 and mc^2 in (22) and (24). Thus

*Using the PV relations $G = e_*^2/m_*^2$ and $e_*^2 = r_* m_* c^2$ in the curvature force leads to $mc^2/r = mm_*G/rr_*$ and shows the direct gravitational interaction between the electron mass and the Planck particle masses within the PV.

both the Pemper derivation and the Dirac equation argue compellingly for the existence of the Planck vacuum state and its place in the physical scheme of things. It is noted in passing that the force in (24) vanishes at the electron's Compton radius $r_c = e_*^2/mc^2$.

Appendix A: Galilean Coordinate System

The laboratory system in which the charge propagates is considered to be a Galilean reference system. In that system (x, y, z) represents the radius vector from the system origin to any field point (considered in the calculations to be fixed). The position of the charge traveling at a constant rate v along the positive z-axis is $(0, 0, vt)$; so at time $t = 0$ the charge crosses the origin. Since the field point is fixed, the vector in the x-y plane

$$\mathbf{b} = b\,\widehat{\mathbf{b}} \equiv \mathbf{x} + \mathbf{y} \tag{A1}$$

is constant. The radius vector from the position of the charge to the field point is then

$$\mathbf{r} = (x, y, z - vt)\,. \tag{A2}$$

Combining (A1) and (A2) gives

$$r = \left[b^2 + (z - vt)^2\right]^{1/2} \tag{A3}$$

for the magnitude of that vector.

If θ is the angle between the radius $\mathbf{r}$ and the positive z-axis, it is easy to show from (A1)—(A3) that

$$r \sin\theta = b \tag{A4}$$

and

$$r \cos\theta = z - vt \tag{A5}$$

and from (A3)—(A5) that

$$\dot{r} = -v \cos\theta \tag{A6}$$

and

$$r\dot{\theta} = v \sin\theta, \tag{A7}$$

where the overhead dot represents a partial derivative with respect to time.

From (7) the initial magnetic field in the charge-PV interaction is

$$B_1 = \beta E_0 \sin\theta = \beta \cdot \frac{e}{r^2} \cdot \frac{b}{r} = \frac{\beta e b}{\left[b^2 + (z - vt)^2\right]^{3/2}} \tag{A8}$$

whose time differential leads to

$$\dot{B}_1 = \frac{3c\beta^2 E_0 \sin\theta\cos\theta}{r} \tag{A9}$$

in a straightforward manner.
From (11) in the text

$$B_2 = \beta E_1 \sin\theta = \frac{3\beta^3 e b^3}{2\left[b^2 + (z - vt)^2\right]^{5/2}} - \lambda_1 B_1, \tag{A10}$$

which leads to

$$\dot{B}_2 = \frac{15c\beta^4 E_o \sin^3\theta\cos\theta}{2r} - \lambda_1 \dot{B}_1\,. \tag{A11}$$

From $B_3 = \beta E_2 \sin\theta$,

$$\begin{aligned} B_3 &= \frac{15\beta^5 E_0 \sin^5\theta}{8} - \lambda_1 \frac{3\beta^3 E_0 \sin^3\theta}{2} - \lambda_2 \beta E_0 \sin\theta \\ &= \frac{15\beta^5 e b^5}{8\left[b^2 + (z - vt)^2\right]^{7/2}} - \lambda_1 \frac{3\beta^3 e b^3}{2\left[b^2 + (z - vt)^2\right]^{5/2}} \\ &\quad -\lambda_2 \frac{\beta e b}{\left[b^2 + (z - vt)^2\right]^{3/2}} \end{aligned} \tag{A12}$$

and

$$\begin{aligned} \dot{B}_3 &= \frac{3 \cdot 5 \cdot 7c\beta^6 E_0 \sin^5\theta\cos\theta}{8r} - \lambda_1 \frac{3 \cdot 5c\beta^4 E_0 \sin^3\theta\cos\theta}{2r} \\ &\quad -\lambda_2 \frac{3c\beta^2 E_0 \sin\theta\cos\theta}{r}\,. \end{aligned} \tag{A13}$$

Appendix B: Lorentz Transformed Fields

The Lorentz transformation coefficients $a_{\mu\nu}$ in the coordinate transformation [7, pp. 380–381]

$$\begin{aligned} x'_\mu = a_{\mu\nu} x_\mu &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & i\beta\gamma \\ 0 & 0 & -i\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ ict \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \\ \gamma(z - vt) \\ i\gamma(ct - \beta z) \end{pmatrix} \end{aligned} \tag{B1}$$

lead to the Lorentz transformed fields

$$F'_{\mu\nu} = a_{\mu\sigma} a_{\nu\tau} F_{\sigma\tau}, \tag{B2}$$

where the $F'_{\mu\nu}$, etc., are the electromagnetic field tensors. The primed and unprimed parameters refer respectively to the charge-at-rest and laboratory systems, where the charge system travels along the z-axis of the laboratory system with a constant velocity v.

Using the static Coulomb field in the charge system and transforming it to the laboratory system with the inverse of (B2) leads to the magnitude

$$E = \frac{\gamma e\left[b^2 + (z - vt)^2\right]^{1/2}}{\left[b^2 + \gamma(z - vt)^2\right]^{3/2}} \tag{B3}$$

for the electric field, where $\gamma = 1/(1 - \beta^2)^{1/2}$. (B3) reduces to (19) in the following way:

$$E = \frac{\gamma e\left[b^2 + (z - vt)^2\right]^{1/2}}{\gamma^3\left[b^2 + (z - vt)^2 - \beta^2 b^2\right]^{3/2}}$$

$$= \frac{e/\left[b^2 + (z - vt)^2\right]}{\gamma^2 \left[1 - \beta^2 b^2/[b^2 + (z - vt)^2]\right]^{3/2}}$$

$$= \frac{(1 - \beta^2)\, E_0}{\left(1 - \beta^2 \sin^2 \theta\right)^{3/2}} \,. \tag{B4}$$

Submitted on January 5, 2011 / Accepted on January 6, 2011

References

1. Pemper R.R. A classical foundation for electrodynamics. Master Dissertation, Univ. of Texas, El Paso, 1977. Barnes T.G. Physics of the Future – A Classical Unification of Physics, Institute for Creation Research, California, 1983.
2. Daywitt W.C. The Planck vacuum. *Progress in Physics*, 2009, v. 1, 20–26.
3. Daywitt W.C. The Source of the Quantum Vacuum. *Progress in Physics*, 2009, v. 1, 27–32.
4. Gingrich D.M. Practical Quantum Electrodynamics, CRC – The Taylor & Francis Group, Boca Raton, 2006.
5. Daywitt W.C. The Dirac Electron in the Planck Vacuum Theory. *Progress in Physics*, 2010, v. 4, 69–71.
6. Puthoff H.E. Gravity as a zero-point-fluctuation force. *Physical Review A*, 1989, v. 39, no. 5, 2333–2342.
7. Jackson J.D. Classical Electrodynamics. John Wiley & Sons, 1st ed., 2nd printing, New York, 1962.

Charged Polaritons with Spin 1

Vahan Minasyan and Valentin Samoilov
Scientific Center of Applied Research, JINR, Joliot-Curie 6, Dubna, 141980, Russia
E-mails: mvahan@scar.jinr.ru; scar@off-serv.jinr.ru

We present a new model for metal which is based on the stimulated vibration of independent charged Fermi-ions, representing as independent harmonic oscillators with natural frequencies, under action of longitudinal and transverse elastic waves. Due to application of the elastic wave-particle principle and ion-wave dualities, we predict the existence of two types of charged Polaritons with spin 1 which are induced by longitudinal and transverse elastic fields. As result of presented theory, at small wavenumbers, these charged polaritons represent charged phonons.

1 Introduction

In our recent paper [1], we proposed a new model for dielectric materials consisting of neutral Fermi atoms. By the stimulated vibration of independent charged Fermi-atoms, representing as independent harmonic oscillators with natural frequencies by actions of the longitudinal and transverse elastic waves, due to application of the principle of elastic wave-particle duality, we predicted the lattice of a solid consists of two types of Sound Boson-Particles with spin 1, with finite masses around 500 times smaller than the atom mass. Namely, we had shown that these lattice Sound-Particles excite the longitudinal and transverse phonons with spin 1. In this context, we proposed new model for solids representing as dielectric substance which is different from the well-known models of Einstein [2] and Debye [3] because: 1), we suggest that the atoms are the Fermi particles which are absent in the Einstein and Debye models; 2), we consider the stimulated oscillation of atoms by action of longitudinal and transverse lattice waves which in turn consist of the Sound Particles.

Thus, the elastic lattice waves stimulate the vibration of the fermion-atoms with one natural wavelength, we suggested that ions have two independent natural frequencies by under action of a longitudinal and a transverse wave. Introduction of the application of the principle of elastic wave-particle duality as well as the model of hard spheres we found an appearance of a cut off in the spectrum energy of phonons which have spin 1 [1].

In this letter, we treat the thermodynamic property of metal under action of the ultrasonic waves. We propose a new model for metal where the charged Fermi-ions vibrate with natural frequencies Ω_l and Ω_t, by under action of longitudinal and transverse elastic waves. Thus, we consider a model for metal as independent charged Fermi-ions of lattice and gas of free electrons or free Frölich-Schafroth charged bosons (singlet electron pairs) [4]. Each charged ion is coupled with a point of lattice knot by spring, creating an ion dipole [5,6]. The lattice knots define the equilibrium positions of all ions which vibrate with natural frequencies Ω_l and Ω_t, under action of longitudinal and transverse elastic fields which in turn leads to creation of the transverse electromagnetic fields moving with speeds c_l and c_t. These transverse electromagnetic waves describe the ions by the principle of ion-wave duality [7]. Using the representation of the electromagnetic field structure of one ion with ion-wave duality in analogous manner, as it was presented in a homogenous medium for an electromagnetic wave [8], we obtain that the neutral phonons cannot be excited in such substances as metals, they may be induced only in dielectric material [1]. In this respect, we find the charged polaritons with spin 1 which are always excited in a metal, and at small wavenumbers, they represent as charged phonons.

2 New model for metal

The Einstein model of a solid considers the solid as gas of N atoms in a box with volume V. Each atom is coupled with a point of the lattice knot. The lattice knots define the dynamical equilibrium position of each atom which vibrates with natural frequency Ω_0. The vibration of atom occurs near equilibrium position corresponding to the minimum of potential energy (harmonic approximation of close neighbors). We presented the model of ion-dipoles [5,6] which represents ions coupled with points of lattice knots. It differs from the Einstein model of solids where the neutral independent atoms are considered in lattice knots, these ions are vibrating with natural frequencies Ω_l and Ω_t forming ion-dipoles by under action longitudinal and transverse ultrasonic lattice fields.

Usually, matters are simplified assuming the transfer of heat from one part of the body to another occurs very slowly. This is a reason to suggest that the heat exchange during times of the order of the period of oscillatory motions in the body is negligible, therefore, we can regard any part of the body as thermally insulated, and there occur adiabatic deformations. Since all deformations are supposed to be small, the motions considered in the theory of elasticity are small elastic oscillations. In this respect, the equation of motion for elastic continuum medium [9] represents as

$$\varrho\ddot{\vec{u}} = c_t^2\nabla^2\vec{u} + (c_l^2 - c_t^2)\ grad\ div\ \vec{u}, \qquad (1)$$

where $\vec{u} = \vec{u}(\vec{r}, t)$ is the vectorial displacement of any particle in the solid; c_l and c_t are, respectively, the velocities of a longitudinal and a transverse ultrasonic wave.

We shall begin by discussing a plane longitudinal elastic wave with condition $curl\ \vec{u} = 0$ and a plane transverse elastic wave with condition $div\ \vec{u} = 0$ in an infinite isotropic medium. In this respect, the vector displacement $\vec{u}$ is the sum of the vector displacements of a longitudinal u_l and of a transverse ultrasonic wave u_t:

$$\vec{u} = \vec{u}_l + \vec{u}_t. \tag{2}$$

In turn, the equations of motion for a longitudinal and a transverse elastic wave take the form of the wave-equations:

$$\nabla^2 \vec{u}_l - \frac{1}{c_l^2} \frac{d^2 \vec{u}_l}{dt^2} = 0, \tag{3}$$

$$\nabla^2 \vec{u}_t - \frac{1}{c_t^2} \frac{d^2 \vec{u}_l}{dt^2} = 0. \tag{4}$$

It is well known, in quantum mechanics, a matter wave is determined by electromagnetic wave-particle duality or de Broglie wave of matter [7]. We argue that in analogous manner, we may apply the elastic wave-particle duality. This reasoning allows us to present a model of elastic field as the Bose-gas consisting of the Sound Bose-particles with spin 1 having non-zero rest masses which are interacting with each other. In this respect, we may express the vector displacements of a longitudinal u_l and of a transverse ultrasonic wave u_t via the second quantization vector wave functions of Sound Bosons as

$$\vec{u}_l = C_l \left(\vec{\phi}(\vec{r}, t) + \vec{\phi}^{+}(\vec{r}, t) \right) \tag{5}$$

and

$$\vec{u}_t = C_t \left(\vec{\psi}(\vec{r}, t) + \vec{\psi}^{+}(\vec{r}, t) \right), \tag{6}$$

where C_l and C_t are unknown constant normalization coefficients; $\vec{\phi}(\vec{r}, t)$ and $\vec{\phi}^{+}(\vec{r}, t)$ are, respectively, the second quantization wave vector functions for one Sound-Particle, corresponding to the longitudinal elastic wave, at coordinate $\vec{r}$ and time t; $\vec{\psi}(\vec{r}, t)$ and $\vec{\psi}^{+}(\vec{r}, t)$ are, respectively, the second quantization wave vector functions for one Sound-Particle, corresponding to the transverse elastic wave, at coordinate $\vec{r}$ and time t:

$$\vec{\phi}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k},\sigma} \vec{a}_{\vec{k},\sigma} e^{i(\vec{k}\vec{r} + kc_l t)} \tag{7}$$

$$\vec{\phi}^{+}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k},\sigma} \vec{a}^{+}_{\vec{k},\sigma} e^{-i(\vec{k}\vec{r} + kc_l t)} \tag{8}$$

and

$$\vec{\psi}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k},\sigma} \vec{b}_{\vec{k},\sigma} e^{i(\vec{k}\vec{r} + kc_t t)} \tag{9}$$

$$\vec{\psi}^{+}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k},\sigma} \vec{b}^{+}_{\vec{k},\sigma} e^{-i(\vec{k}\vec{r} + kc_t t)}, \tag{10}$$

where $\vec{a}^{+}_{\vec{k},\sigma}$ and $\vec{a}_{\vec{k},\sigma}$ are, respectively, the Bose vector-operators of creation and annihilation for one free longitudinal Sound Particle with spin 1, described by a vector $\vec{k}$ whose direction gives the direction of motion of the longitudinal wave; $\vec{b}^{+}_{\vec{k},\sigma}$ and $\vec{b}_{\vec{k},\sigma}$ are, respectively, the Bose vector-operators of creation and annihilation for one free transverse Sound Particle with spin 1, described by a vector $\vec{k}$ whose direction gives the direction of motion of the transverse wave.

In this respect, the vector-operators $\vec{a}^{+}_{\vec{k},\sigma}, \vec{a}_{\vec{k},\sigma}$ and $\vec{b}^{+}_{\vec{k},\sigma}, \vec{b}_{\vec{k},\sigma}$ satisfy the Bose commutation relations as:

$$\left[\hat{a}_{\vec{k},\sigma}, \hat{a}^{+}_{\vec{k}',\sigma'} \right] = \delta_{\vec{k},\vec{k}'} \cdot \delta_{\sigma,\sigma'}$$

$$[\hat{a}_{\vec{k},\sigma}, \hat{a}_{\vec{k}',\sigma'}] = 0$$

$$[\hat{a}^{+}_{\vec{k},\sigma}, \hat{a}^{+}_{\vec{k}',\sigma'}] = 0$$

and

$$\left[\hat{b}_{\vec{k},\sigma}, \hat{b}^{+}_{\vec{k}',\sigma'} \right] = \delta_{\vec{k},\vec{k}'} \cdot \delta_{\sigma,\sigma'}$$

$$[\hat{b}_{\vec{k},\sigma}, \hat{b}_{\vec{k}',\sigma'}] = 0$$

$$[\hat{b}^{+}_{\vec{k},\sigma}, \hat{b}^{+}_{\vec{k}',\sigma'}] = 0.$$

Thus, as we see the vector displacements of a longitudinal u_l and of a transverse ultrasonic wave u_t satisfy the wave-equations of (3) and (4) because they have the following forms due to application of (5) and (6):

$$\vec{u}_l = \frac{C_l}{\sqrt{V}} \sum_{\vec{k},\sigma} \left(\vec{a}_{\vec{k},\sigma} e^{i(\vec{k}\vec{r} + kc_l t)} + \vec{a}^{+}_{\vec{k},\sigma} e^{-i(\vec{k}\vec{r} + kc_l t)} \right) \tag{11}$$

and

$$\vec{u}_t = \frac{C_t}{\sqrt{V}} \sum_{\vec{k},\sigma} \left(\vec{b}_{\vec{k},\sigma} e^{i(\vec{k}\vec{r} + kc_t t)} + \vec{b}^{+}_{\vec{k},\sigma} e^{-i(\vec{k}\vec{r} + kc_t t)} \right). \tag{12}$$

In this context, we may emphasize that the Bose vector operators $\vec{a}^{+}_{\vec{k},\sigma}$, $\vec{a}_{\vec{k},\sigma}$ and $\vec{b}^{+}_{\vec{k}}, \sigma$ and $\vec{b}_{\vec{k}}, \sigma$ communicate with each other because the vector displacements of a longitudinal $\vec{u}_l$ and a transverse ultrasonic wave $\vec{u}_t$ are independent, and in turn, satisfy the condition of a scalar multiplication $\vec{u}_l \cdot \vec{u}_t = 0$.

Consequently, the Hamiltonian operator $\hat{H}$ of the system, consisting of the vibrating Fermi-ions with mass M, is represented in the following form:

$$\hat{H} = \hat{H}_l + \hat{H}_t, \tag{13}$$

where

$$\hat{H}_l = \frac{MN}{V}\int\left(\frac{d\vec{u}_l}{dt}\right)^2 dV + \frac{NM\Omega_l^2}{V}\int(\vec{u}_l)^2 dV \tag{14}$$

and

$$\hat{H}_t = \frac{MN}{V}\int\left(\frac{d\vec{u}_t}{dt}\right)^2 dV + \frac{NM\Omega_t^2}{V}\int(\vec{u}_t)^2 dV, \tag{15}$$

where Ω_l and Ω_t are, respectively, the natural frequencies of the atom through action of the longitudinal and transverse elastic waves.

To find the Hamiltonian operator $\hat{H}$ of the system, we use the formalism of Dirac [10]:

$$\frac{d\vec{u}_l}{dt} = \frac{ic_lC_l}{\sqrt{V}}\sum_{\vec{k},\sigma} k\left(\vec{a}_{\vec{k},\sigma}e^{ikc_lt} - \vec{a}^{+}_{-\vec{k},\sigma}e^{-ikc_lt}\right)e^{i\vec{k}\vec{r}} \tag{16}$$

and

$$\frac{d\vec{u}_t}{dt} = \frac{ic_tC_t}{\sqrt{V}}\sum_{\vec{k},\sigma} k\left(\vec{b}_{\vec{k},\sigma}e^{ikc_tt} - \vec{b}^{+}_{-\vec{k},\sigma}e^{-ikc_tt}\right)e^{i\vec{k}\vec{r}}, \tag{17}$$

which by substituting into (14) and (15), using (11) and (12), gives the reduced form of the Hamiltonian operators $\hat{H}_l$ and $\hat{H}_t$:

$$\hat{H}_l = \sum_{\vec{k},\sigma}\left(\frac{2MNC_l^2c_l^2k^2}{V} + \frac{2MNC_l^2\Omega_l^2}{V}\right)\vec{a}^{+}_{\vec{k},\sigma}\vec{a}_{\vec{k},\sigma} -$$

$$- \sum_{\vec{k},\sigma}\left(\frac{2MNC_l^2c_l^2k^2}{V} - \frac{2MNC_l^2\Omega_l^2}{V}\right)\left(a_{\vec{k},\sigma}\vec{a}_{-\vec{k},\sigma} + a^{+}_{-\vec{k}}\vec{a}^{+}_{\vec{k},\sigma}\right) \tag{18}$$

and

$$\hat{H}_t = \sum_{\vec{k},\sigma}\left(\frac{2MNC_t^2c_t^2k^2}{V} + \frac{2MNC_t^2\Omega_t^2}{V}\right)\vec{b}^{+}_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma} -$$

$$- \sum_{\vec{k},\sigma}\left(\frac{2MNC_t^2c_t^2k^2}{V} - \frac{2MNC_t^2\Omega_t^2}{V}\right)\left(b_{\vec{k},\sigma}\vec{b}_{-\vec{k},\sigma} + b^{+}_{-\vec{k}}\vec{b}^{+}_{\vec{k},\sigma}\right), \tag{19}$$

where the normalization coefficients C_l and C_t are defined by the first term of right side of (18) and (19) which represent the kinetic energies of longitudinal Sound Particles $\frac{\hbar^2k^2}{2m_l}$ and transverse Sound Particles $\frac{\hbar^2k^2}{2m_t}$ with masses m_l and m_t, respectively. Therefore we suggest to find C_l and C_t:

$$\frac{2MNC_l^2c_l^2k^2}{V} = \frac{\hbar^2k^2}{2m_l} \tag{20}$$

and

$$\frac{2MNC_t^2c_t^2k^2}{V} = \frac{\hbar^2k^2}{2m_t}, \tag{21}$$

which in turn determine

$$C_l = \frac{\hbar}{2c_l\sqrt{m_l\rho}} \tag{22}$$

and

$$C_t = \frac{\hbar}{2c_t\sqrt{m_t\rho}}, \tag{23}$$

where $\rho = \frac{MN}{V}$ is the density of solid.

As we had shown in [1], at absolute zero $T = 0$, the Fermi ions fill the Fermi sphere in momentum space. Thus, there are two type Fermi atoms by the value of its spin z-component $\mu = \pm\frac{1}{2}$ with the boundary wave number k_f of the Fermi, which, in turn, is determined by a condition:

$$\frac{V}{2\pi^2}\int_0^{k_f} k^2dk = \frac{N}{2},$$

where N is the total number of Fermi-ions in the solid. This reasoning together with the model of hard spheres claims the important condition to introduce the boundary wave number $k_f = \left(\frac{3\pi^2N}{V}\right)^{\frac{1}{3}}$ coinciding with k_l and k_t. Then, there is an important condition $k_f = k_l = k_t$ which determines a relationship between natural oscillator frequencies

$$k_f = \frac{\Omega_l}{c_l} = \frac{\Omega_t}{c_t}. \tag{24}$$

3 Charged Polaritons

In papers [5, 6], we demostrated the so-called transformation of longitudinal and transverse elastic waves into transverse electromagnetic fields with vectors of the electric waves $\vec{E}_l$ and $\vec{E}_t$, corresponding to the ion displacements $\vec{u}_l$ and $\vec{u}_t$, respectively. In turn, the equations of motion are presented in the following forms [5, 6]:

$$M\frac{d^2\vec{u}_l}{dt^2} + M\Omega_l^2\vec{u}_l = -e\vec{E}_l \tag{25}$$

and

$$M\frac{d^2\vec{u}_t}{dt^2} + M\Omega_t^2\vec{u}_t = -e\vec{E}_t. \tag{26}$$

The vector of the electric waves $\vec{E}_l$ and $\vec{E}_t$ are defined by substitution of the meaning of $\vec{u}_l$ and $\vec{u}_t$ from (11) and (12), respectively, into (25) and (26):

$$\vec{E}_l(\vec{r},t)=\frac{C_l}{e\sqrt{V}}\sum_{\vec{k},\sigma}\gamma_{\vec{k},l}\left(\vec{a}_{\vec{k},\sigma}e^{i(\vec{k}\vec{r}+kc_lt)}+\vec{a}^{+}_{\vec{k},\sigma}e^{-i(\vec{k}\vec{r}+kc_lt)}\right) \quad (27)$$

and

$$\vec{E}_t(\vec{r},t)=\frac{C_t}{e\sqrt{V}}\sum_{\vec{k},\sigma}\gamma_{\vec{k},t}\left(\vec{b}_{\vec{k},\sigma}e^{i(\vec{k}\vec{r}+kc_tt)}+\vec{b}^{+}_{\vec{k},\sigma}e^{-i(\vec{k}\vec{r}+kc_tt)}\right), \quad (28)$$

where

$$\gamma_{\vec{k},l}=M\left(k^2c_l^2-\Omega_l^2\right) \quad (29)$$

and

$$\gamma_{\vec{k},t}=M\left(k^2c_t^2-\Omega_t^2\right). \quad (30)$$

On the other hand, by action of the longitudinal and transverse ultrasonic waves on the charged ion [5, 6], these ultrasonic waves are transformed into transverse electromagnetic fields with electric wave vectors $\vec{E}_l$ and $\vec{E}_t$ which in turn describe the de Broglie wave of charged ions expressed via electric $\vec{E}_l(\vec{r},t)$ and $\vec{E}_t(\vec{r},t)$ fields of one ion-wave particle in homogeneous medium. In fact, these electric $\vec{E}_l(\vec{r},t)$ and $\vec{E}_t(\vec{r},t)$ fields satisfy the Maxwell's equations in dielectric medium:

$$curl\,\vec{H}_l-\frac{\varepsilon_l}{c}\frac{d\vec{E}_l}{dt}=0 \quad (31)$$

$$curl\,\vec{E}_l+\frac{1}{c}\frac{d\vec{H}_l}{dt}=0 \quad (32)$$

$$div\,\vec{E}_l=0 \quad (33)$$

$$div\,\vec{H}_l=0 \quad (34)$$

and

$$curl\,\vec{H}_t-\frac{\varepsilon_t}{c}\frac{d\vec{E}_t}{dt}=0 \quad (35)$$

$$curl\,\vec{E}_t+\frac{1}{c}\frac{d\vec{H}_t}{dt}=0 \quad (36)$$

$$div\,\vec{E}_t=0 \quad (37)$$

$$div\,\vec{H}_t=0 \quad (38)$$

with

$$\sqrt{\varepsilon_l}=\frac{c}{c_l} \quad (39)$$

and

$$\sqrt{\varepsilon_l}=\frac{c}{c_l}, \quad (40)$$

where $\vec{H}_l=\vec{H}_l(\vec{r},t)$ and $\vec{H}_t=\vec{H}_t(\vec{r},t)$ are, respectively, the local magnetic fields, corresponding to longitudinal and transverse ultrasonic waves, depending on space coordinate $\vec{r}$ and time t; ε_l and ε_t are, respectively, the dielectric constants for transverse electric fields $\vec{E}_l(\vec{r},t)$ and $\vec{E}_t(\vec{r},t)$ corresponding to longitudinal and transverse ultrasonic waves; c is the velocity of electromagnetic wave in vacuum; $\mu=1$ is the magnetic susceptibility.

When using Eqs. (31–40) and results of letter [8], we may present the transverse electric fields $\vec{E}_l(\vec{r},t)$ and $\vec{E}_t(\vec{r},t)$ by the quantization forms:

$$\vec{E}_l(\vec{r},t)=\frac{A_l}{\sqrt{V}}\sum_{\vec{k}}\left(\vec{c}_{\vec{k}}e^{i(\vec{k}\vec{r}+kc_lt)}+\vec{c}^{+}_{\vec{k}}e^{-i(\vec{k}\vec{r}+kc_lt)}\right) \quad (41)$$

and

$$\vec{E}_t(\vec{r},t)=\frac{A_t}{\sqrt{V}}\sum_{\vec{k}\neq0}\left(\vec{d}_{\vec{k}}e^{i(\vec{k}\vec{r}+kc_tt)}+\vec{d}^{+}_{\vec{k}}e^{-i(\vec{k}\vec{r}+kc_tt)}\right), \quad (42)$$

where A_l and A_t are the unknown constants which are found as below; $\vec{c}^{+}_{\vec{k}}$, $\vec{d}^{+}_{\vec{k}}$ and $\vec{c}_{\vec{k}}$, $\vec{d}_{\vec{k}}$ are, respectively, the Bose vector-operators of creation and annihilation of electric fields of one ion-wave particle with wave vector $\vec{k}$ which are directed along of the wave normal $\vec{s}$ or $\vec{k}=k\vec{s}$. These Bose vector-operators $\vec{E}_l(\vec{r},t)$ and $\vec{E}_t(\vec{r},t)$ are directed to the direction of the unit vectors $\vec{l}$ and $\vec{t}$ which are perpendicular to the wave normal $\vec{s}$; $\hat{N}$ is the operator total number of charged ions.

In this context, we indicate that the vector-operators $\vec{c}^{+}_{\vec{k},\sigma}$, $\vec{c}_{\vec{k},\sigma}$ and $\vec{d}^{+}_{\vec{k},\sigma}$, $\vec{d}_{\vec{k},\sigma}$ satisfy the Bose commutation relations as:

$$\left[\hat{c}_{\vec{k},\sigma},\hat{c}^{+}_{\vec{k}',\sigma'}\right]=\delta_{\vec{k},\vec{k}'}\cdot\delta_{\sigma,\sigma'}$$

$$[\hat{c}_{\vec{k},\sigma},\hat{c}_{\vec{k}',\sigma'}]=0$$

$$[\hat{c}^{+}_{\vec{k},\sigma},\hat{c}^{+}_{\vec{k}',\sigma'}]=0$$

and

$$\left[\hat{d}_{\vec{k},\sigma},\hat{d}^{+}_{\vec{k}',\sigma'}\right]=\delta_{\vec{k},\vec{k}'}\cdot\delta_{\sigma,\sigma'}$$

$$[\hat{d}_{\vec{k},\sigma},\hat{d}_{\vec{k}',\sigma'}]=0$$

$$[\hat{d}^{+}_{\vec{k},\sigma},\hat{d}^{+}_{\vec{k}',\sigma'}]=0.$$

Comparing (41) with (27) and (42) with (28), we get

$$\vec{d}_{\vec{k},\sigma}=\frac{eA_l}{C_l\gamma_{\vec{k},l}}\vec{c}_{\vec{k},\sigma} \quad (43)$$

and

$$\vec{b}_{\vec{k},\sigma} = \frac{eA_t}{C_t\gamma_{\vec{k},t}}\vec{d}_{\vec{k},\sigma}. \tag{44}$$

Now, substituting $\vec{a}_{\vec{k},\sigma}$ and $\vec{b}_{\vec{k},\sigma}$ into (18) and (19), we obtain the reduced form of the Hamiltonian operators $\hat{H}_l$ and $\hat{H}_l$ which are expressed via terms of the electric fields of the ion-wave particle:

$$\begin{aligned}\hat{H}_l = \sum_{\vec{k},\sigma}\frac{e^2A_l^2}{\gamma^2_{\vec{k},l}}\left[\left(\frac{2MNc_l^2k^2}{V}+\frac{2MN\Omega_l^2}{V}\right)\vec{c}^{+}_{\vec{k},\sigma}c_{\vec{k},\sigma}-\right.\\ \left.-\left(\frac{MNc_l^2k^2}{V}-\frac{MN\Omega_l^2}{V}\right)\left(\vec{c}_{-\vec{k},\sigma}\vec{c}_{\vec{k},\sigma}+\vec{c}^{+}_{\vec{k},\sigma}\vec{c}^{+}_{-\vec{k},\sigma}\right)\right]\end{aligned} \tag{45}$$

and

$$\begin{aligned}\hat{H}_t = \sum_{\vec{k},\sigma}\frac{e^2A_t^2}{\gamma^2_{\vec{k},l}}\left[\left(\frac{2MNc_t^2k^2}{V}+\frac{2MN\Omega_t^2}{V}\right)\vec{d}^{+}_{\vec{k},\sigma}d_{\vec{k},\sigma}-\right.\\ \left.-\left(\frac{MNc_t^2k^2}{V}-\frac{MN\Omega_t^2}{V}\right)\left(\vec{d}_{-\vec{k},\sigma}\vec{d}_{\vec{k},\sigma}+\vec{d}^{+}_{\vec{k},\sigma}\vec{d}^{+}_{-\vec{k},\sigma}\right)\right].\end{aligned} \tag{46}$$

To evaluate the energy levels of the operators $\hat{H}_l$ (45) and $\hat{H}_t$ (46) within the diagonal form, we use a transformation of the vector-Bose-operators:

$$\vec{c}_{\vec{k},\sigma} = \frac{\vec{l}_{\vec{k},\sigma}+L_{\vec{k}}\vec{l}^{+}_{-\vec{k},\sigma}}{\sqrt{1-L^2_{\vec{k}}}} \tag{47}$$

and

$$\vec{d}_{\vec{k},\sigma} = \frac{\vec{t}_{\vec{k},\sigma}+M_{\vec{k}}\vec{t}^{+}_{-\vec{k},\sigma}}{\sqrt{1-M^2_{\vec{k}}}}, \tag{48}$$

where $L_{\vec{k}}$ and $M_{\vec{k}}$ are, respectively, the real symmetrical functions of a wave vector $\vec{k}$.

Consequently,

$$\hat{H}_l = \sum_{k<k_f,\sigma}\varepsilon_{\vec{k},l}\vec{l}^{+}_{\vec{k},\sigma}\vec{l}_{\vec{k},\sigma} \tag{49}$$

and

$$\hat{H}_t = \sum_{k<k_f,\sigma}\varepsilon_{\vec{k},t}\vec{t}^{+}_{\vec{k},\sigma}\vec{t}_{\vec{k},\sigma} \tag{50}$$

at

$$L^2_{\vec{k}} = \frac{\frac{2MNc_l^2k^2}{V}+\frac{2MN\Omega_l^2}{V}-\varepsilon_{\vec{k},l}}{\frac{2MNc_l^2k^2}{V}+\frac{2MN\Omega_l^2}{V}+\varepsilon_{\vec{k},l}}$$

$$M^2_{\vec{k}} = \frac{\frac{2MNc_t^2k^2}{V}+\frac{2MN\Omega_l^2}{V}-\varepsilon_{\vec{k},t}}{\frac{2MNc_t^2k^2}{V}+\frac{2MN\Omega_l^2}{V}+\varepsilon_{\vec{k},t}}.$$

Hence, we infer that the Bose-operators $\vec{l}^{+}_{\vec{k},\sigma}$, $\vec{l}_{\vec{k},\sigma}$ and $\vec{t}^{+}_{\vec{k},\sigma}$, $\vec{t}_{\vec{k},\sigma}$ are, respectively, the vector of "creation" and the vector of "annihilation" operators of charged polaritons with spin 1 with the energies:

$$\varepsilon_{\vec{k},l} = \frac{4e^2\rho c_lA_l^2\Omega_lk}{\gamma^2_{\vec{k},l}} \tag{51}$$

and

$$\varepsilon_{\vec{k},t} = \frac{4e^2\rho c_tA_t^2\Omega_tk}{\gamma^2_{\vec{k},t}}. \tag{52}$$

Hence, we note that these polaritons are charged because the Hamiltonian contains the square of charge, e^2. This picture is similar to the Coulomb interaction between two charges.

Obviously, at small wave numbers $k \ll \frac{\Omega_l}{c_l}$ and $k \ll \frac{\Omega_t}{c_t}$, these charged polaritons are presented as charged phonons with energies:

$$\varepsilon_{\vec{k},l} \approx \hbar kv_l \tag{53}$$

and

$$\varepsilon_{\vec{k},t} \approx \hbar kv_t, \tag{54}$$

where $v_l = \frac{4\rho c_le^2A_l^2}{\hbar M^2\Omega_l^3}$ and $v_t = \frac{4\rho c_te^2A_t^2}{\hbar M^2\Omega_t^3}$ are, respectively, the velocities of charged phonons with spin 1 corresponding to the longitudinal and transverse acoustic fields. To find the unknown constants A_l^2 and A_t^2, we suggest that $v_l = c_l$ and $v_t = c_t$ as it was presented in [1]. This suggestion leads to the results obtained in [1] and in turn presented in Debye's theory. Thus, when choosing $A_l^2 = \frac{\hbar M^2\Omega_l^3}{4\rho e^2}$ and $A_t^2 = \frac{\hbar M^2\Omega_t^3}{4\rho e^2}$, the energies of charged polaritons represent as

$$\varepsilon_{\vec{k},l} = \frac{\hbar\Omega_l^4c_lk}{\left(k^2c_l^2-\Omega_l^2\right)^2} \tag{55}$$

and

$$\varepsilon_{\vec{k},t} = \frac{\hbar\Omega_t^4c_tk}{\left(k^2c_t^2-\Omega_t^2\right)^2}, \tag{56}$$

which at large wave numbers $k \gg \frac{\Omega_l}{c_l}$ and $k \gg \frac{\Omega_t}{c_t}$, and taking into account (24), leads to the following form for the energies of charged polaritons:

$$\varepsilon_{\vec{k},l} = \frac{\hbar k_f^4c_l}{k^3}. \tag{57}$$

In fact, the stimulated vibration of ions by elastic waves lead to the formation of the charged polaritons with spin 1.

Thus, we predicted the existence of a new type of charged quasiparticles in nature. On the other hand, we note that the quantization of elastic fields is fulfilled for the new model of metals. In analogous manner, as it was presented in [1], we may show that the acoustic field operator does not commute with its momentum density.

Submitted on January 6, 2011 / Accepted on January 10, 2011

References

1. Minasyan V. N., Samoilov V. N. Sound-Particles and Phonons with Spin 1. *Progress in Physics*, 2011, v. 1, 81–86.
2. Einstein A. Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Warme. *Annalen der Physik*, 1907, v. 22, 180–190.
3. Debye P. Zur Theorie der spezifischen Waerme. *Annalen der Physik*, 1912, v. 39, 789–839.
4. Minasyan V. N., Samoilov V. N. Formation of Singlet Fermion Pairs in the Dilute Gas of Boson-Fermion Mixture. *Progress in Physics*, 2010, v. 4, 3–9.
5. Minasyan V. N., Samoilov V. N. The Intensity of the Light Diffraction by Supersonic Longitudinal Waves in Solid, *Progress in Physics*, 2010, v. 2, 60–63.
6. Minasyan V. N., Samoilov V. N. Dispersion of Own Frequency of Ion-dipole by Supersonic Transverse Wave in Solid, *Progress in Physics*, 2010, v. 4, 10–12.
7. de Broglie L. Researches on the quantum theory. *Annalen der Physik*, 1925, v. 3, 22–32.
8. Minasyan V. N., Samoilov V. N. New resonance-polariton Bose-quasiparticles enhances optical transmission into nanoholes in metal films. *Physics Letters A*, 2011, v. 375, 698–711.
9. Landau L. D., Lifshiz E. M. Theory of Elasticity. *Theoretical Physics*, 1987, v. 11, 124–127.
10. Dirac P. A. M. The Principles of Quantum Mechanics. Clarendon press, Oxford, (1958).

New Fundamental *Light Particle* and Breakdown of Stefan-Boltzmann's Law

Vahan Minasyan and Valentin Samoilov
Scientific Center of Applied Research, JINR, Dubna, 141980, Russia
E-mails: mvahan@scar.jinr.ru; scar@off-serv.jinr.ru

Recently, we predicted the existence of fundamental particles in Nature, neutral *Light Particles* with spin 1 and rest mass $m = 1.8{\times}10^{-4}m_e$, in addition to electrons, neutrons and protons. We call these particles Light Bosons because they create electromagnetic field which represents Planck's gas of massless photons together with a gas of *Light Particles* in the condensate. Such reasoning leads to a breakdown of Stefan–Boltzmann's law at low temperature. On the other hand, the existence of new fundamental neutral *Light Particles* leads to correction of such physical concepts as Bose-Einstein condensation of photons, polaritons and exciton polaritons.

1 Introduction

First, the quantization scheme for the local electromagnetic field in vacuum was presented by Planck in his black body radiation studies [1]. In this context, the classical Maxwell equations lead to appearance of the so-called ultraviolet catastrophe; to remove this problem, Planck proposed the model of the electromagnetic field as an ideal Bose gas of massless photons with spin one. However, Dirac [2] showed the Planck photon-gas could be obtained through a quantization scheme for the local electromagnetic field, presenting a theoretical description of the quantization of the local electromagnetic field in vacuum by use of a model Bose-gas of local plane electromagnetic waves propagating by speed c in vacuum.

In a different way, in regard to Plank and Dirac's models, we consider the structure of the electromagnetic field [3] as a non-ideal gas consisting of N neutral *Light Bose Particles* with spin 1 and finite mass m, confined in a box of volume V. The form of potential interaction between *Light Particles* is defined by introduction of the principle of wave-particle duality of de Broglie [4] and principle of gauge invariance. In this respect, a non-ideal Bose-gas consisting of *Light Particles* with spin 1 and non-zero rest mass is described by Planck's gas of massless photons together with a gas consisting of *Light Particles* in the condensate. In this context, we defined the *Light Particle* by the model of hard sphere particles [5]. Such definition of *Light Particles* leads to cutting off the spectrum of the electromagnetic wave by the boundary wave number $k_0 = \frac{mc}{\hbar}$ or boundary frequency $\omega_\gamma = 10^{18}$ Hz of gamma radiation at the value of the rest mass of the *Light Particle* $m = 1.8 \times 10^{-4}m_e$. On the other hand, the existence of the boundary wave number $k_0 = \frac{mc}{\hbar}$ for the electromagnetic field in vacuum is connected with the characteristic length of the interaction between two neighboring *Light Bosons* in the coordinate space with the minimal distance $d = \frac{1}{k_0} = \frac{\hbar}{mc} = 2{\times}10^{-9}m$. This reasoning determines the density of *Light Bosons* $\frac{N}{V}$ as $\frac{N}{V} = \frac{3}{4\pi d^3} = 0.3{\times}10^{26}m^{-3}$.

It is well known that Stefan-Boltzmann's law [6] for thermal radiation, presented by Planck's formula [1], determines the average energy density $\frac{U}{V}$ as

$$\frac{U}{V} = \frac{2}{V}\sum_{0\le k<\infty} \hbar kc\overline{i^{+}_{\vec{k}} i_{\vec{k}}} = \sigma T^4, \qquad (1)$$

where $\hbar$ is the Planck constant; σ is the Stefan-Boltzmann constant; $\overline{i^{+}_{\vec{k}} i_{\vec{k}}}$ is the average number of photons with the wave vector $\vec{k}$ at the temperature T:

$$\overline{i^{+}_{\vec{k}} i_{\vec{k}}} = \frac{1}{e^{\frac{\hbar kc}{kT}} - 1}. \qquad (2)$$

Obviously, at $T = 0$, the average energy density vanishes in Eq.(1), i.e. $\frac{U}{V} = 0$, which follows from Stefan-Boltzmann's law.

However, as we show, the existence of the predicted *Light Particles* breaks Stefan-Boltzmann's law for black body radiation at low temperature.

2 Breakdown of Stefan-Boltzmann's law

Now, we consider the results of letter [3], where the average energy density of black radiation $\frac{U}{V}$ is represented as:

$$\frac{U}{V} = \frac{mc^2 N_{0,T}}{V} + \frac{2}{V}\sum_{0\le k<k_0} \hbar kc\overline{i^{+}_{\vec{k}} i_{\vec{k}}}, \qquad (3)$$

where $\frac{mc^2 N_{0,T}}{V}$ is a new term, in regard to Plank's formula (1), which determines the energy density of *Light Particles* in the condensate; $\frac{N_{0,T}}{V}$ is the density of *Light Particles* in the condensate.

In this respect, the equation for the density of *Light Particles* in the condensate $\frac{N_{0,T}}{V}$ represents as

$$\frac{N_{0,T}}{V} = \frac{N}{V} - \frac{1}{V}\sum_{0<k<k_0} \frac{L^2_{\vec{k}}}{1 - L^2_{\vec{k}}} - \frac{1}{V}\sum_{0<k<k_0} \frac{1 + L^2_{\vec{k}}}{1 - L^2_{\vec{k}}}\overline{i^{+}_{\vec{k}} i_{\vec{k}}} \qquad (4)$$

with the real symmetrical function $L_{\vec{k}}$ from the wave vector $\vec{k}$:

$$L^2_{\vec{k}} = \frac{\frac{\hbar^2 k^2}{2m} + \frac{mc^2}{2} - \hbar kc}{\frac{\hbar^2 k^2}{2m} + \frac{mc^2}{2} + \hbar kc}. \quad (5)$$

Our calculation shows that at absolute zero the value of $\overline{\vec{i}^+_{\vec{k}} \vec{i}_{\vec{k}}} = 0$, and therefore the average energy density of black radiation $\frac{U}{V}$ reduces to

$$\frac{U}{V} = \frac{mc^2 N_{0,T=0}}{V} = \frac{mc^2 N}{V} - \frac{m^4 c^5 B(2,3)}{4\pi^2 \hbar^3} \approx \frac{mc^2 N}{V}, \quad (6)$$

where $B(2,3) = \int_0^1 x(1-x)^2 dx = 0.1$ is the beta function.

Thus, the average energy density of black radiation $\frac{U}{V}$ is a constant at absolute zero. In fact, there is a breakdown of Stefan-Boltzmann's law for thermal radiation.

In conclusion, it should be also noted that *Light Bosons* in vacuum create photons, while *Light Bosons* in a homogeneous medium generate the so-called polaritons. This fact implies that photons and polaritons are quasiparticles, therefore, Bose-Einstein condensation of photons [7], polaritons [8] and exciton polaritons [9] has no physical sense.

Acknowledgements

This work is dedicated to the memory of the Great British Physicist Prof. Marshall Stoneham, F.R.S., (London Centre for Nanotechnology, and Department of Physics and Astronomy University College London, UK), who helped us with English. We are very grateful to him.

Submitted on January 6, 2011 / Accepted on January 25, 2011

References

1. Planck M. On the Law of Distribution of Energy in the Normal Spectrum. *Annalen der Physik*, 1901, v. 4, 553–563.
2. Dirac P. A. M. The Principles of Quantum Mechanics. Clarendon Press, Oxford, 1958.
3. Minasyan V. N., Samoilov V. N. New resonance-polariton Bose-quasiparticles enhances optical transmission into nanoholes in metal films. *Physics Letters A*, 2011, v. 375, 698–711.
4. de Broglie L. Researches on the quantum theory. *Annalen der Physik*, 1925, v. 3, 22–32.
5. Huang K. Statistical Mechanics. John Wiley, New York, 1963.
6. Stefan J. Uber die Beziehung zwischen der Warmestrahlung und der Temperatur. In: Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften, Bd. 79 (Wien 1879), 391–428.
7. Klaers J., Schmitt J., Vewinger F., Weitz M. Bose-Einstein condensation of photons in an optical microcavity. *Nature*, 2010, v. 468, 545–548.
8. Balili R., Hartwell V., Snoke D., Pfeiffer L., West K. Bose-Einstein condensation of microcavity polaritons in a trap. *Science*, 2007, v. 316, 1007–1010.
9. Kasprzak J. et al. Bose-Einstein condensation of exciton polaritons. *Nature*, 2006, v. 443, 409–414.

The Point Mass Concept

Bo Lehnert

Alfvén Laboratory, Royal Institute of Technology, SE–10044 Stockholm, Sweden. E-mail: Bo.Lehnert@ee.kth.se

A point-mass concept has been elaborated from the equations of the gravitational field. One application of these deductions results in a black hole configuration of the Schwarzschild type, having no electric charge and no angular momentum. The critical mass of a gravitational collapse with respect to the nuclear binding energy is found to be in the range of 0.4 to 90 solar masses. A second application is connected with the speculation about an extended symmetric law of gravitation, based on the options of positive and negative mass for a particle at given positive energy. This would make masses of equal polarity attract each other, while masses of opposite polarity repel each other. Matter and antimatter are further proposed to be associated with the states of positive and negative mass. Under fully symmetric conditions this could provide a mechanism for the separation of antimatter from matter at an early stage of the universe.

1 Introduction

In connection with an earlier elaborated revised quantum-electrodynamic theory, a revised renormalisation procedure has been developed to solve the problem of infinite self-energy of the point-charge-like electron [1, 2]. In the present investigation an analogous procedure is applied to the basic equations of gravitation, to formulate a corresponding point mass concept. Two applications result from such a treatment. The first concerns the special Schwarzschild case of a black hole with its critical limit of gravitational collapse. The second application is represented by the speculation about an extended form of the gravitation law, in which full symmetry is obtained by including both positive and negative mass concepts. This further leads to the question whether such concepts could have their correspondence in matter and antimatter, and in their mutual separation.

2 The conventional law of gravitation

In this investigation the analysis is limited to the steady case of spherical symmetry, in a corresponding frame where r is the only independent variable.

2.1 Basic equations

Following Bergmann [3], a steady gravitational field strength

$$\mathbf{g} = -\nabla\phi \tag{1}$$

is considered which originates from the potential $\phi(r)$. The source of the field strength is a mass density ρ related to $\mathbf{g}$ by

$$-\mathrm{div}\,\mathbf{g} = 4\pi G\rho = \nabla^2\phi = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\phi}{dr}\right), \tag{2}$$

where $G = 6.6726\times10^{-11}\ \mathrm{m^3kg^{-1}s^{-2}}$ is the constant of gravitation in SI units. The associated force density becomes

$$\mathbf{f} = \rho\,\mathbf{g}\,. \tag{3}$$

In the conventional interpretation there only exists a positive mass density $\rho \geqslant 0$. This makes in a way the gravitational field asymmetric, as compared to the electrostatic field which includes both polarities of electric charge density.

A complete form of the potential ϕ would consist of a series of both positive and negative powers of r, but the present analysis will be restricted and simplified by studying each power separately, in the form

$$\phi(r) = \phi_0\left(\frac{r}{r_0}\right)^{\alpha}. \tag{4}$$

Here ϕ_0 is a constant, r_0 represents a characteristic dimension and α is a positive or negative integer. Equation (2) yields

$$4\pi G\rho = \frac{\phi_0}{r_0^{\alpha}}\,\alpha(\alpha+1)\,r^{\alpha-2} > 0\,. \tag{5}$$

When limiting the investigations by the condition $\rho > 0$, the cases $\alpha = 0$ and $\alpha = -1$ have to be excluded, leaving the regimes of positive $\alpha = (1, 2, \ldots)$ and negative $\alpha = (-2, -3, \ldots)$ to be considered for *positive* values of ϕ_0.

2.2 Point mass formation

For reasons to become clear from the deductions which follow, we now study a spherical configuration in which the mass density ρ is zero within an inner hollow region $0 \leqslant r \leqslant r_i$, and where $\rho > 0$ in the outer region $r > r_i$. From relation (5) the total integrated mass $P(r)$ inside the radius r then becomes

$$P(r) = \int_0^r \rho\,4\pi r^2 dr = \frac{1}{G}\frac{\phi_0}{r_0^{\alpha}}\,\alpha\left(r^{\alpha+1} - r_i^{\alpha+1}\right) > 0 \tag{6}$$

with a resulting local field strength $\mathbf{g} = (g, 0, 0)$ given by

$$g(r) = -G\frac{P(r)}{r^2} = -\frac{\phi_0}{r_0^{\alpha}}\,\frac{\alpha}{r^2}\left(r^{\alpha+1} - r_i^{\alpha+1}\right) < 0 \tag{7}$$

and a local force density $\mathbf{f} = (f, 0, 0)$ where

$$f(r) = -\frac{1}{4\pi G}\frac{\phi_0}{r_0^{\alpha}}^2 \alpha^2(\alpha+1) r^{\alpha-4}\left(r^{\alpha+1} - r_i^{\alpha+1}\right) < 0\,. \quad (8)$$

Here all (P, g, f) refer to the range $r \geqslant r_i$, and $\phi_0 > 0$.

A distinction is further made between the two regimes of positive and negative α:

- When $\alpha = (1, 2, \ldots)$ of a convergent potential (4), this hollow configuration has an integrated mass (6) which increases monotonically with r, from zero at $r = r_i$ to large values. This behaviour is the same for a vanishing r_i and does not lead to a point-like mass at small r_i.
- When $\alpha = (-2, -3, \ldots)$ of a divergent potential (4), the hollow configuration leads to a point-mass-like geometry at small r_i. This is similar to a point-charge-like geometry earlier treated in a model of the electron [1, 2], and will be considered in the following analysis.

2.3 The renormalised point mass

In the range $\alpha \leqslant -2$ expressions (6)–(8) are preferably cast into a form with $\gamma = -\alpha \geqslant 2$ where

$$P(r) = \frac{1}{G}\left(\phi_0 r_0^{\gamma}\right)\gamma\left(r_i^{-\gamma+1} - r^{-\gamma+1}\right) > 0\,, \quad (9)$$

$$g(r) = -\left(\phi_0 r_0^{\gamma}\right)\frac{\gamma}{r^2}\left(r_i^{-\gamma+1} - r^{-\gamma+1}\right) < 0\,, \quad (10)$$

$$f(r) = -\frac{1}{4\pi G}\left(\phi_0 r_0^{\gamma}\right)^2 \gamma^2(\gamma-1) r^{-\gamma-4}\left(r_i^{-\gamma+1} - r^{-\gamma+1}\right) < 0. \quad (11)$$

Here an erroneous result would be obtained if the terms including r_i are dropped and the hollow configuration is abandoned. Due to eqs. (9)–(11) this would namely result in a negative mass P, a positive field strength g, and a repulsive local gravitational force density f.

The radius r_i of the hollow inner region is now made to approach zero. The total integrated mass of eq. (9) is then concentrated to an infinitesimally small layer. Applying a revised renormalisation procedure in analogy with an earlier scheme [1, 2], we "shrink" the combined parameters $\phi_0 r_0^{\gamma}$ and $r_i^{\gamma-1}$ in such a way that

$$\phi_0 r_0^{\gamma} = c_{\phi r}\cdot\varepsilon \qquad r_i^{\gamma-1} = c_i\cdot\varepsilon \qquad 0 < \varepsilon \ll 1\,, \quad (12)$$

where ε is a smallness parameter and $c_{\phi r}$ and c_i are positive constants. A further introduction of

$$P_0 \equiv \frac{1}{G}\frac{\gamma\, c_{\phi r}}{c_i} \quad (13)$$

results in $P(r) = 0$ for $r \leqslant r_i$ and

$$P(r) = P_0\left[1 - \left(\frac{r_i}{r}\right)^{\gamma-1}\right] \qquad r > r_i\,. \quad (14)$$

In the limit $\varepsilon \to 0$ and $r_i \to 0$ there is then a point mass P_0 at the origin. This mass generates a field strength

$$g(r) = -G\frac{P_0}{r^2} \quad (15)$$

at the distance r according to equations (10), (12) and (13). With another point mass P_1 at the distance r, there is a mutual attraction force

$$F_{01} = P_1 g(r) = -G\frac{P_0 P_1}{r^2}\,, \quad (16)$$

which is identical with the gravitation law for two point masses.

To further elucidate the result of eqs. (12)-(16) it is first observed that, in the conventional renormalisation procedure, the divergent behaviour of an infinite self-energy is outbalanced by adding extra infinite ad-hoc *counter-terms* to the Lagrangian, to obtain a finite difference between two "infinities". Even if such a procedure has been successful, however, it does not appear to be quite acceptable from the logical and physical points of view. The present revised procedure represented by expressions (12) implies on the other hand that the "infinity" of the divergent potential ϕ_0 at a shrinking radius r_i is instead outbalanced by the "zeros" of the inherent shrinking *counter-factors* $c_{\phi r}\cdot\varepsilon$ and $c_i\cdot\varepsilon$.

3 A black hole of Schwarzschild type

A star which collapses into a black hole under the compressive action of its own gravitational field is a subject of ever increasing interest. In its most generalized form the physics of the black hole includes both gravitational and electromagnetic fields as well as problems of General Relativity, to account for its mass, net electric charge, and its intrinsic angular momentum. The associated theoretical analysis and related astronomical observations have been extensively described in a review by Misner, Thorne and Wheeler [4] among others. Here the analysis of the previous section will be applied to the far more simplified special case by Schwarzschild, in which there is no electric charge and no angular momentum. Thereby it has also to be noticed that no black hole in the universe has a substantial electric charge [4].

3.1 The inward directed gravitational pressure

From eq. (11) is seen that the inward directed local force density is zero for $r \leqslant r_i$, increases with r to a maximum within a thin shell, and finally drops to zero at large r. The integrated inward directed gravitational pressure on this shell thus becomes

$$p = \int_0^{\infty} f\,dr = -\frac{G}{8\pi}\frac{(\gamma-1)^2}{(\gamma+1)(\gamma+3)}\frac{P_0^2}{r_i^4} \quad (17)$$

in the limit of small ε and r_i.

When this pressure becomes comparable to the relevant energy density of the compressed matter, a corresponding gravitational collapse is expected to occur.

3.2 Gravitational collapse of the nuclear binding forces

Here we consider the limit at which matter is compressed into a body of densely packed nucleons, and when the pressure of eq. (17) tends to exceed the energy density of the nucleon binding energy. The radius of a nucleus is [5]

$$r_N = 1.5 \times 10^{-15} A^{1/2} \quad [\mathrm{m}], \tag{18}$$

where A is the mass number. A densely packed sphere of N nuclei has the volume

$$V_N = \frac{4}{3}\pi N r_N^3 = \frac{4}{3}\pi r_{eq}^3 , \tag{19}$$

where r_{eq} is the equivalent radius of the sphere. The total binding energy of a nucleus is further conceived as the work required to completely dissociate it into its component nucleons. This energy is about 8 MeV per nucleon [5, 6]. With A nucleons per nucleus, the total binding energy of a body of N nuclei thus becomes

$$W_N = NA\, w_N , \tag{20}$$

where $w_N = 8$ MeV $= 1.28 \times 10^{-12}$ J. The equivalent binding energy density of the body is then

$$p_N = \frac{W_N}{V_N} = \frac{3}{4}\frac{A w_N}{\pi r_N^3} = 0.907 \times 10^{32} A^{-1/2} \quad [\mathrm{J \times m^{-3}}]. \tag{21}$$

The shell-like region of gravitational pressure has a force density (11) which reaches its maximum at the radius

$$r_m = r_i \left(\frac{2\gamma + 3}{\gamma + 4}\right)^{1/(\gamma-1)} \tag{22}$$

being only a little larger than r_i. This implies that the radius r_{eq} of eq. (19) is roughly equal to r_i and

$$r_i \cong r_N N^{1/3}. \tag{23}$$

With N nuclei of the mass Am_p and m_p as the proton mass, the total mass becomes

$$P_0 = NAm_p , \tag{24}$$

which yields

$$r_i \cong r_N \left(\frac{P_0}{Am_p}\right)^{1/3} = 1.26 \times 10^{-6} A^{1/6} P_0^{1/3} \quad [\mathrm{m}]. \tag{25}$$

This result finally combines with eq. (17) to an equivalent gravitational pressure

$$p = -1.1 \times 10^{12} \frac{(\gamma-1)^2}{(\gamma+1)(\gamma+3)} A^{-2/3} P_0^{2/3} \quad [\mathrm{J \times m^{-3}}]. \tag{26}$$

For a gravitational collapse defined by $-p > p_N$ the point mass P_0 then has to exceed the critical limit

$$P_{0c} \cong 7.5 \times 10^{29} \left[\frac{(\gamma+1)(\gamma+3)}{(\gamma-1)^2}\right]^{3/2} A^{1/4} \quad [\mathrm{kg}]. \tag{27}$$

For $\gamma \geqslant 2$ and $1 \leqslant A \leqslant 250$, the critical mass would then be found in the range of about $0.4 \leqslant P_{0c} \leqslant 90$ solar masses of about 1.98×10^{30} kg.

4 Speculations about a generalized law of gravitation

The Coulomb law of interaction between electrically charged bodies is symmetric in the sense that it includes both polarities of charge and attractive as well as repulsive forces. The classical Newtonian law of gravitation includes on the other hand only one polarity of mass and only attractive forces. In fact, this asymmetry does not come out as a necessity from the basic equations (1)–(3) of a curl-free gravitational field strength. The question could therefore be raised whether a more general and symmetric law of gravitation could be deduced from the same equations, and whether this could have a relevant physical interpretation.

4.1 Mass polarity

In relativistic mechanics the momentum $\mathbf{p}$ of a particle with the velocity $\mathbf{u}$ and rest mass m_0 becomes [7]

$$\mathbf{p} = m_0 \mathbf{u} \left[1 - \left(\frac{u}{c}\right)^2\right]^{-1/2} . \tag{28}$$

With the energy E of the particle, the Lorentz invariance further leads to the relation

$$p^2 - \frac{E^2}{c^2} = -m_0^2 c^2, \tag{29}$$

where $p^2 = \mathbf{p}^2$ and $u^2 = \mathbf{u}^2$. Equations (28) and (29) yield

$$E^2 = m_0^2 c^4 \left[1 - \left(\frac{u}{c}\right)^2\right]^{-1} \equiv m^2 c^4, \tag{30}$$

leading in principle to two roots

$$E = \pm mc^2. \tag{31}$$

In this investigation the discussion is limited to a positive energy E, resulting in positive and negative gravitational masses

$$m = \pm \frac{E}{c^2} \qquad E > 0 . \tag{32}$$

This interpretation differs from that of the negative energy states of positrons proposed in the "hole" theory by Dirac [8] corresponding to the plus sign in eq. (31) and where both E and m are negative.

4.2 An extended law of gravitation

With the possibility of negative gravitational masses in mind, we now return to the potential ϕ of equations (1) and (4) where the amplitude factor ϕ_0 can now adopt both positive and negative values, as defined by the notation $\phi_{0+} > 0$ and $\phi_{0-} < 0$, and where corresponding subscripts are introduced for (P, g, f, P_0) of eqs. (9)–(11) and (13). Then $P_+ > 0$, $P_- < 0$, $P_{0+} > 0$, $P_{0-} < 0$, $g_+ < 0$, $g_- > 0$, but $f_+ < 0$ and $f_- < 0$ always represent an attraction force due to the quadratic dependence on ϕ_0 in eq. (11). With P_{1+} or P_{1-} as an additional point mass

at the distance r from P_{0+} or P_{0-}, there is then an extended form of the law (16), as represented by the forces

$$P_{1+}g_{+} = -G\frac{P_{0+}P_{1+}}{r^2} = P_{1-}g_{-} = -G\frac{P_{0-}P_{1-}}{r^2}\,, \qquad (33)$$

$$P_{1+}g_{-} = G\frac{P_{0-}P_{1+}}{r^2} = P_{1-}g_{+} = G\frac{P_{0-}P_{1+}}{r^2}\,. \qquad (34)$$

These relations are symmetric in the gravitational force interactions, where masses of equal polarity attract each other, and masses of opposite polarity repel each other. It would imply that the interactions in a universe consisting entirely of negative masses would become the same as those in a universe consisting entirely of positive masses. In this way specific mass polarity could, in fact, become a matter of definition.

4.3 A possible rôle of antimatter

At this point the further question may be raised whether the states of positive and negative mass could be associated with those of matter and antimatter, respectively. A number of points become related to such a proposal.

The first point concerns an experimental test of the repulsive behaviour due to eq. (34). If an electrically neutral beam of anti-matter, such as of antihydrogen atoms, could be formed in a horizontal direction, such a beam would be deflected upwards if consisting of negative mass. However, the deflection is expected to be small and difficult to measure.

A model has earlier been elaborated for a particle with elementary charge, being symmetric in its applications to the electron and the positron [1, 2]. The model includes an electric charge q_0, a rest mass m_0, and an angular momentum s_0 of the particle. The corresponding relations between included parameters are easily seen to be consistent with electron-positron pair formation in which $q_0 = -e$, $m_0 = +E/c^2$ and $s_0 = +h/4\pi$ for the electron and $q_0 = +e$, $m_0 = -E/c^2$ and $s_0 = +h/4\pi$ for the positron when the formation is due to a photon of spin $+h/2\pi$. The energy of the photon is then at least equal to $2E$ where the electron and the photon both have positive energies E.

The energy of photons and their electromagnetic radiation field also have to be regarded as an equivalent mass due to Einstein's mass-energy relation. This raises the additional question whether full symmetry also requires the photon to have a positive or negative gravitational mass, as given by

$$m_\nu = \pm\frac{h\nu}{c^2}\,. \qquad (35)$$

If equal proportions of matter and antimatter would have been formed at an early stage of the universe, the repulsive gravitational force between their positive and negative masses could provide a mechanism which expels antimatter from matter and vice versa, also under fully symmetric conditions. Such a mechanism can become important even if the gravitational forces are much weaker than the electrostatic ones, because matter and antimatter are expected to appear as electrically quasi-neutral cosmical plasmas. The final result would come out to be separate universes of matter and antimatter.

In a theory on the metagalaxy, Alfvén and Klein [9] have earlier suggested that there should exist limited regions in our universe which contain matter or antimatter, and being separated by thin boundary layers within which annihilation reactions take place. A simplified model of such layers has been established in which the matter-antimatter "ambiplasma" is immersed in a unidirectional magnetic field [10]. The separation of the cells of matter from those of antimatter by a confining magnetic field geometry in three spatial directions is, however, a problem of at least the same complication as that of a magnetically confined fusion reactor.

5 Conclusions

From the conventional equations of the gravitational field, the point-mass concept has in this investigation been elaborated in terms of a revised renormalisation procedure. In a first application a black hole configuration of the Schwarzschild type has been studied, in which there is no electric charge and no angular momentum. A gravitational collapse in respect to the nuclear binding energy is then found to occur at a critical point mass in the range of about 0.4 to 90 solar masses. This result becomes modified if the collapse is related to other restrictions such as to the formation of "primordial black holes" growing by the accretion of radiation and matter [4], or to phenomena such as a strong centrifugal force.

A second application is represented by the speculation about an extended law of gravitation, based on the options of positive and negative mass of a particle at a given positive energy, and on the basic equations for a curl-free gravitational fieldstrength. This would lead to a fully symmetric law due to which masses of equal polarity attract each other, and masses of opposite polarity repel each other. A further proposal is made to associate matter and antimatter with the states of positive and negative mass. Even under fully symmetric conditions, this provides a mechanism for separating antimatter from matter at an early stage of development of the universe.

After the completion of this work, the author has been informed of a hypothesis with negative mass by Choi [11], having some points in common with the present paper.

Submitted on January 15, 2011 / Accepted on January 17, 2011

References

1. Lehnert B. A revised electromagnetic theory with fundamental applications. Swedish Physics Archive. Edited by D. Rabounski, The National Library of Sweden, Stockholm, 2008; and Bogolyubov Institute for Theoretical Physics. Edited by Z. I. Vakhnenko and A. Zagorodny, Kiev, 2008.
2. Lehnert B. Steady particle states of revised electromagnetics. *Progress in Physics*, 2006, v. 3, 43–50.
3. Bergmann P. G. Introduction to the theory of relativity. Prentice Hall, Inc., New York, 1942, Ch. X.

4. Misner C. W., Thorne K. S., Wheeler J. A. Gravitation. W. H. Freeman and Co., San Francisco and Reading, 1973, Ch. 33.
5. Bethe H. A. Elementary nuclear theory. John Wiley & Sons, Inc., New York, Chapman and Hall, Ltd., London, 1947, pp. 8 and 17.
6. Fermi E. Nuclear physics. Revised edition. The University of Chicago Press, 1950, p. 3.
7. Møller C. The theory of relativity. Oxford, Claredon Press, 1952, Ch. III.
8. Schiff L. I. Quantum mechanics. McGrawHill Book Comp., Inc., New York, Toronto, London, 1949, Ch. XII, Sec. 44.
9. Alfvén H., Klein O. Matter-antimatter annihilation and cosmology. *Arkiv för Fysik*, 1962, v. 23, 187–195.
10. Lehnert B. Problems of matter-antimatter boundary layers. *Astrophysics and Space Science*, 1977, v. 46, 61–71.
11. Choi Hyoyoung. Hypothesis of dark matter and dark energy with negative mass. viXra.org/abs/0907.0015, 2010.

Quark Annihilation and Lepton Formation versus Pair Production and Neutrino Oscillation: The Fourth Generation of Leptons

T. X. Zhang

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

The emergence or formation of leptons from particles composed of quarks is still remained very poorly understood. In this paper, we propose that leptons are formed by quark-antiquark annihilations. There are two types of quark-antiquark annihilations. Type-I quark-antiquark annihilation annihilates only color charges, which is an incomplete annihilation and forms structureless and colorless but electrically charged leptons such as electron, muon, and tau particles. Type-II quark-antiquark annihilation annihilates both electric and color charges, which is a complete annihilation and forms structureless, colorless, and electrically neutral leptons such as electron, muon, and tau neutrinos. Analyzing these two types of annihilations between up and down quarks and antiquarks with an excited quantum state for each of them, we predict the fourth generation of leptons named lambda particle and neutrino. On the contrary quark-antiquark annihilation, a lepton particle or neutrino, when it collides, can be disintegrated into a quark-antiquark pair. The disintegrated quark-antiquark pair, if it is excited and/or changed in flavor during the collision, will annihilate into another type of lepton particle or neutrino. This quark-antiquark annihilation and pair production scenario provides unique understanding for the formation of leptons, predicts the fourth generation of leptons, and explains the oscillation of neutrinos without hurting the standard model of particle physics. With this scenario, we can understand the recent OPERA measurement of a tau particle in a muon neutrino beam as well as the early measurements of muon particles in electron neutrino beams.

1 Introduction

Elementary particles can be categorized into hadrons and leptons in accord with whether they participate in the strong interaction or not. Hadrons participate in the strong interaction, while leptons do not. All hadrons are composites of quarks [1-3]. There are six types of quarks denoted as six different flavors: up, down, charm, strange, top, and bottom, which are usually grouped into three generations: $\{u, d\}$, $\{c, s\}$, $\{t, b\}$. Color charge is a fundamental property of quarks, which has analogies with the notion of electric charge of particles. There are three varieties of color charges: red, green, and blue. An antiquark's color is antired, antigreen, or antiblue. Quarks and antiquarks also hold electric charges but they are fractional, $\pm e/3$ or $\pm 2e/3$, where $e = 1.6 \times 10^{-19}$ C is the charge of proton.

There are also six types of leptons discovered so far, which are electron, muon, and tau particles and their corresponding neutrinos. These six types of leptons are also grouped into three generations: $\{e^-, \nu_e\}$, $\{\mu^-, \nu_\mu\}$, $\{\tau^-, \nu_\tau\}$. The antiparticles of the charged leptons have positive charges. It is inappropriate to correspond the three generations of leptons to the three generations of quarks because all these three generations of leptons are formed or produced directly in association with only the first generation of quarks. We are still unsure that how leptons form and whether the fourth generation of leptons exists or not [4-8].

In this paper, we propose that leptons, including the fourth generation, are formed by quark-antiquark annihilations. Electrically charged leptons are formed when the color charges of quarks and antiquarks with different flavors are annihilated, while neutrinos are formed when both the electric and color charges of quarks and antiquarks with the same flavor are annihilated. We also suggest that quarks and antiquarks can be produced in pairs from disintegrations of leptons. This quark-antiquark annihilation and pair production model predicts the fourth generation of leptons and explains the measurements of neutrino oscillations.

2 Quark Annihilation and Lepton Formation

Quark-antiquark annihilation is widely interested in particle physics [9-13]. A quark and an antiquark may annihilate to form a lepton. There are two possible types of quark-antiquark annihilations. Type-I quark-antiquark annihilation only annihilates their color charges. It is an incomplete annihilation usually occurred between different flavor quark and antiquark and forms structureless and colorless but electrically charged leptons such as e^-, e^+ , μ^-, μ^+, τ^-, and τ^+. Type-II quark-antiquark annihilation annihilates both electric and color charges. It is a complete annihilation usually occurred between same flavor quark and antiquark and forms

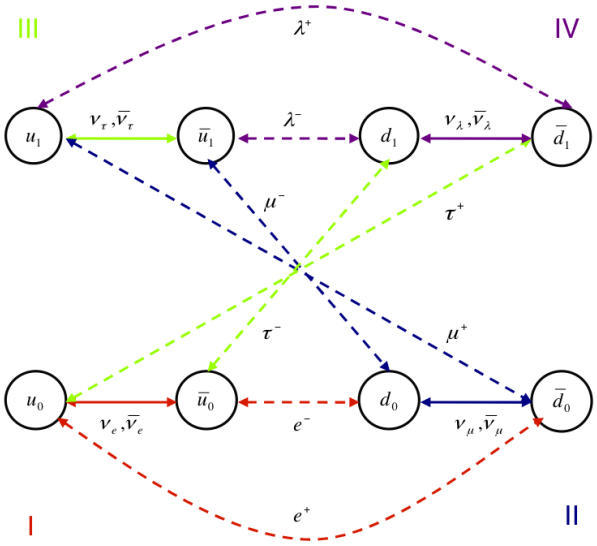


Fig. 1: Formation of the four generations of leptons by annihilations of up and down quarks and antiquarks with an excited quantum state.

Quarks	u_0	d_0	u_1	d_1
$\bar{u}_0$	$\nu_e, \bar{\nu}_e$	e^-	-	τ^-
$\bar{d}_0$	e^+	$\nu_\mu, \bar{\nu}_\mu$	μ^+	-
$\bar{u}_1$	-	μ^-	$\nu_\tau, \bar{\nu}_\tau$	λ^-
$\bar{d}_1$	τ^+	-	λ^+	$\nu_\lambda, \bar{\nu}_\lambda$

Table 1: The up and down quarks and antiquarks in ground and excited quantum states and four generations of leptons

structureless, colorless, and electrically neutral leptons such as ν_e, $\bar{\nu}_e$, ν_μ, $\bar{\nu}_\mu$, ν_τ, and $\bar{\nu}_\tau$.

Mesons are quark-antiquark mixtures without annihilating their charges. For instance, the meson pion π^+ is a mixture of one up quark and one down antiquark. Meson's color charges are not annihilated and thus participate in the strong interaction. Leptons do not participate in the strong interaction because their color charges are annihilated. Particles formed from annihilations do not have structure such as γ-rays formed from particle-antiparticle annihilation. A baryon is a mixture of three quarks such as that a proton is composed of two up quarks and one down quark and that a neutron is composed of one up quark and two down quarks.

Recently, Zhang [14-15] considered the electric and color charges of quarks and antiquarks as two forms of imaginary energy in analogy with mass as a form of real energy and developed a classical unification theory that unifies all natural fundamental interactions with four natural fundamental elements, which are radiation, mass, electric charge, and color charge. According to this consideration, the type-I quark-antiquark annihilation cancels only the color imaginary energies of a quark and a different flavor antiquark, while the type-II quark-antiquark annihilation cancels both the electric and color imaginary energies of a quark and a same flavor antiquark.

Figure 1 is a schematic diagram that shows formations of four generations of leptons from annihilations of up and down quarks and antiquarks with one excited quantum state for each of them. The existence of quark excited states, though not yet directly discovered, has been investigated over three decades [16-18]. That ρ^+ is also a mixture of one up quark and one down antiquark but has more mass than π^+ and many similar examples strongly support that quarks and antiquarks have excited states. In Figure 1, the subscript '0' denotes the ground state and '1' denotes the excited state. The higher excited states are not considered in this study. The dashed arrow lines refer to type-I annihilations of quarks and antiquarks that form electrically charged leptons, while the solid arrow lines refer to type-II annihilations of quarks and antiquarks that form colorless and electrically neutral leptons. These annihilations of quarks and antiquarks and formations of leptons can also be represented in Table 1.

The first generation of leptons is formed by annihilations between the ground state up, ground state antiup, ground state down, and ground state antidown quarks (see the red arrow lines of Figure 1). The up quark u_0 and the antiup quark $\bar{u}_0$ completely annihilate into an electron neutrino ν_e or an electron antineutrino $\bar{\nu}_e$. The antiup quark $\bar{u}_0$ and the down quark d_0 incompletely annihilate into an electron e^-. The up quark u_0 and the antidown quark $\bar{d}_0$ incompletely annihilate into a positron e^+.

The second generation of leptons are formed by annihilations between the ground state down, ground state antidown, excited up, and excited antiup quarks (see the blue arrow lines of Figure 1). The down quark d_0 and the antidown quark $\bar{d}_0$ completely annihilate into a muon neutrino ν_μ or an antimuon neutrino $\bar{\nu}_\mu$. The antiup quark $\bar{u}_1$ and the down quark d_0 incompletely annihilate into a negative muon μ^-. The up quark u_1 and the antidown quark $\bar{d}_0$ incompletely annihilate into a positive muon μ^+.

The third generation of leptons are formed by annihilations between the ground state up, excited up, ground state antiup, excited antiup, excited down, and excited antidown quarks (see the green lines of Figure 1). The up quark u_1 and the antiup quark $\bar{u}_1$ completely annihilate into a tau neutrino ν_τ or a tau antineutrino $\bar{\nu}_\tau$. The antiup quark $\bar{u}_0$ and the down quark d_1 incompletely annihilate into a negative tau τ^-. The up quark u_0 and the antidown quark $\bar{d}_1$ incompletely annihilate into a positive tau τ^+.

The fourth generation of leptons are formed by annihilations between excited up, excited antiup, excited down, and excited antidown quarks (see the purple lines of Figure 1). The down quark d_1 and the antidown quark $\bar{d}_1$ completely annihilate into a lambda neutrino ν_λ or a lambda antineutrino

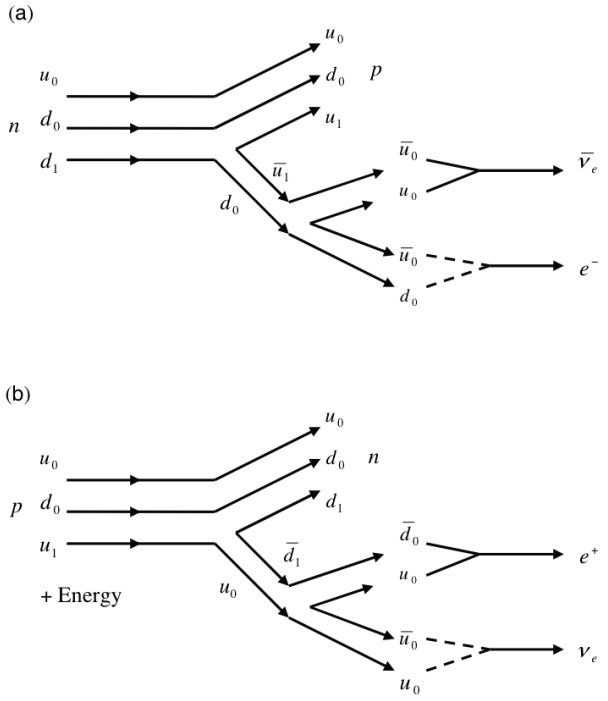


Fig. 2: Formation of the first generation of leptons: (a) e^- and $\bar{\nu}_e$ through beta decay and (b) e^+ and ν_e through positron emission.

$\bar{\nu}_\lambda$. The antiup quark $\bar{u}_1$ and the down quark d_1 incompletely annihilate into a negative lambda λ^-. The up quark u_1 and the antidown quark $\bar{d}_1$ incompletely annihilate into a positive lambda λ^+.

3 Quark Pair Production and Lepton Disintegration

The first generation of leptons can be produced through the beta decay of a neutron, $n \longrightarrow p + e^- + \bar{\nu}_e$ (Figure 2a), and the positron emission of a proton, energy $+ p \longrightarrow n + e^+ + \nu_e$ (Figure 2b).

In the beta decay, an excited down quark in the neutron degenerates into a ground state down quark and an excited up and antiup quark pair, $d_1 \longrightarrow d_0 + (u_1\bar{u}_1)$. The excited antiup quark further degenerates into a ground state up quark and a ground state up and antiup quark pair, $\bar{u}_1 \longrightarrow \bar{u}_0 + (u_0\bar{u}_0)$. The ground state antiup quark incompletely annihilates with the ground state down quark into an electron, $\bar{u}_0 + d_0 \longrightarrow e^-$. The ground state up and antiup quark pair completely annihilates into an electron antineutrino, $u_0 + \bar{u}_0 \longrightarrow \bar{\nu}_e$.

In the positron emission, an excited up quark in the positron after absorbing a certain amount of energy degenerates into a ground state up quark and produces an excited state down and antidown quark pair, energy $+ u_1 \longrightarrow u_0 + (d_1\bar{d}_1)$. The excited antidown quark further degenerates into a ground state antidown quark and produces a ground state up and antiup quark pair, $\bar{d}_1 \longrightarrow \bar{d}_0 + (u_0\bar{u}_0)$. The ground state up quark incompletely annihilates with the ground state antidown quark to form a positron, $u_0 + \bar{d}_0 \longrightarrow e^+$. The ground state up and antiup quark pair completely annihilates into an electron neutrino, $u_0 + \bar{u}_0 \longrightarrow \nu_e$.

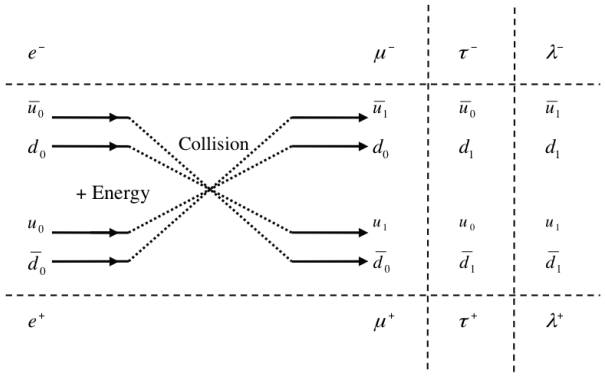


Fig. 3: Production of other three electrically charged leptons from an energetic electron-positron collision. In the collision, electron and positron are first disintegrated into quark-antiquark pairs, which are then excited and annihilated into other generations of electrically charge leptons.

The other three generations of electrically charged leptons can be produced by an energetic electron-positron collision,

$$\text{energy} + e^- + e^+ \longrightarrow \begin{cases} \mu^- + \mu^+ \\ \tau^- + \tau^+ \\ \lambda^- + \lambda^+ \end{cases}, \qquad (1)$$

as also shown in Figure 3. In the particle physics, it has been experimentally shown that the energetic electron-positron collision can produce (μ^-, μ^+) and (τ^-, τ^+). But how the electron-positron collisions produce μ and τ leptons is still remained very poorly understood.

With the quark annihilation and pair production model proposed in this paper, we can understand why an electron-positron can produce μ and τ particles. In addition, we predict the existence of the fourth generation of leptons, λ particle and neutrino. The energetic electron-positron collision disintegrates the electron into a ground state antiup-down quark pair $e^- \longrightarrow (\bar{u}_0 d_0)$ and the positron into a ground state up-antidown quark pair $e^+ \longrightarrow (u_0\bar{d}_0)$. During the collision, the quarks and antiquarks in the disintegrated electron and positron quark-antiquark pairs absorb energy and become excited. The excited quark-antiquark pairs incompletely annihilate into another generation of electrically charged leptons.

There are three possible excitation patterns, which lead to three generations of leptons from the electron-positron collision. If the antiup quark in the disintegrated electron quark-antiquark pair and the up quark in the disintegrated positron quark-antiquark pair are excited, then the annihilations produce leptons μ^- and μ^+. If the down quark in the disintegrated electron quark-antiquark pair and the antidown quark in the disintegrated positron quark-antiquark pair are excited,

then the annihilations produce leptons τ^- and τ^+. If both the antiup and down quarks in the disintegrated electron quark-antiquark pair and both the up and antidown quarks in the disintegrated positron quark-antiquark pair are excited, the annihilations produce the leptons λ^- and λ^+. An electron-positron collision in a different energy level produces a different generation of electrically charged leptons. To produce the λ particles, a more energetic electron-positron collision is required than μ and τ lepton productions. On the other hand, the electron and positron, if they are not disintegrated into quark-antiquark pairs during the collision, can directly annihilate into photons. The disintegrated electron and positron quark-antiquark pairs, if they are excited but not annihilated, can form the weak particles W^- and W^+.

A quark or antiquark can be excited when it absorbs energy or captures a photon. An excited quark or antiquark can degenerate into its corresponding ground state quark or antiquark after it releases a photon and/or one or more quark-antiquark pairs. The decays of these three generations of electrically charged leptons (μ, τ, and λ particles) can produce their corresponding neutrinos through degenerations and annihilations of quarks and antiquarks.

The currently discovered three generations of leptons including the fourth generation predicted in this paper are formed through the annihilations of the up and down quarks and antiquarks with an excited state. All these leptons are corresponding to or associated with the first generation of quarks and antiquarks. Considering the annihilations of other four flavor quarks and antiquarks, we can have many other types of leptons that are corresponding to the second and third generations of quarks and antiquarks. These leptons must be hardly generated and observed because a higher energy is required [4].

4 Quark Annihilation and Pair Production: Neutrino Oscillation

The complete (or type-II) annihilation between a quark and its corresponding antiquark forms a colorless and electrically neutral neutrino. On the contrary quark-antiquark annihilation, a neutrino, when it collides with a nucleon, may be disintegrated into a quark-antiquark pair. The disintegrated quark-antiquark pairs can be excited if it absorbs energy (e.g., $\gamma + u_0 \longrightarrow u_1$) and changed in flavor if it exchanges a weak particle (e.g., $u_0 + W^- \longrightarrow d_0$) during the disintegration. The excited and/or flavor changed quark-antiquark pair then either annihilates into another type of neutrino or interacts with the nucleon to form hadrons and electrically charged leptons. This provides a possible explanation for neutrino oscillations [19-20]. This scenario of neutrino oscillations does not need neutrinos to have mass and thus does not conflict with the standard model of particle physics.

Figure 4 and 5 show all possible oscillations among the four types of neutrinos. An electron neutrino can oscillate into a tau neutrino if the disintegrated quark-antiquark pair ($u_0\bar{u}_0$) is excited into ($u_1\bar{u}_1$) (Figure 4a), a muon neutrino if the disintegrated quark-antiquark pair ($u_0\bar{u}_0$) is changed in flavor into ($d_0\bar{d}_0$) (Figure 4b), and a lambda neutrino if the disintegrated quark-antiquark pair ($u_0\bar{u}_0$) is excited into ($u_1\bar{u}_1$) and then changed in flavor into ($d_1\bar{d}_1$) (Figure 4c). Similarly, A muon neutrino can oscillate into a tau neutrino if the disintegrated quark-antiquark pair ($d_0\bar{d}_0$) is excited and changed in flavor into ($u_1\bar{u}_1$) (Figure 5a) and a lambda neutrino if the disintegrated quark-antiquark pair ($d_0\bar{d}_0$) is excited and changed into ($d_1\bar{d}_1$) (Figure 5b). A tau neutrino can oscillate into a lambda neutrino if the disintegrated quark-antiquark pair ($u_1\bar{u}_1$) is changed in flavor into ($d_1\bar{d}_1$) (Figure 5c). All these oscillations described above are reversible processes. The right arrows in Figures 4 and 5 denote the neutrino oscillations when the disintegrated quark-antiquark pair absorbs energy to be excited or capture weak particles to be changed in flavor. Neutrinos can also oscillate when the disintegrated quark-antiquark pair emits energy and/or releases weak particles. In this cases, the right arrows in Figure 4 and 5 are replaced by left arrows and neutrinos oscillate from heavier ones to lighter ones.

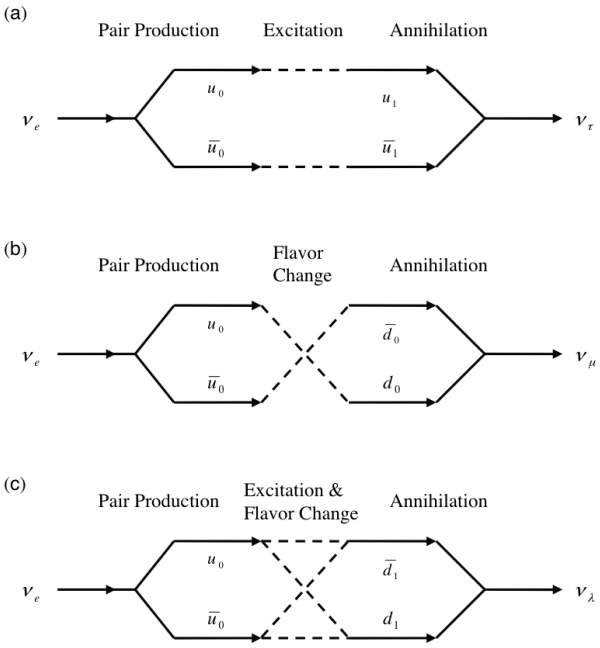


Fig. 4: Neutrino oscillations. (a) Oscillation between electron and tau neutrinos. (b) Oscillation between electron and muon neutrinos. (c) Oscillation between electron and lambda neutrinos.

The recent OPERA experiment at the INFN's Gran Sasso laboratory in Italy first observed directly a tau particle in a muon neutrino beam generated by pion and kaon decays and sent through the Earth from CERN that is 732 km away [21-23]. This significant result can be explained with a muon neutrino disintegration, excitation, and interaction with a nu-

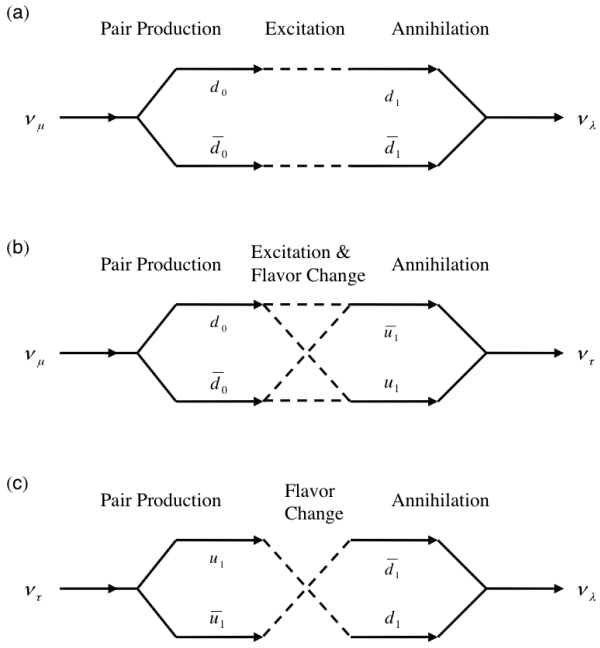


Fig. 5: Neutrino oscillations. (a) Oscillation between muon and lambda neutrinos. (b) Oscillation between muon and tau neutrinos. (c) Oscillation between tau and lambda neutrinos.

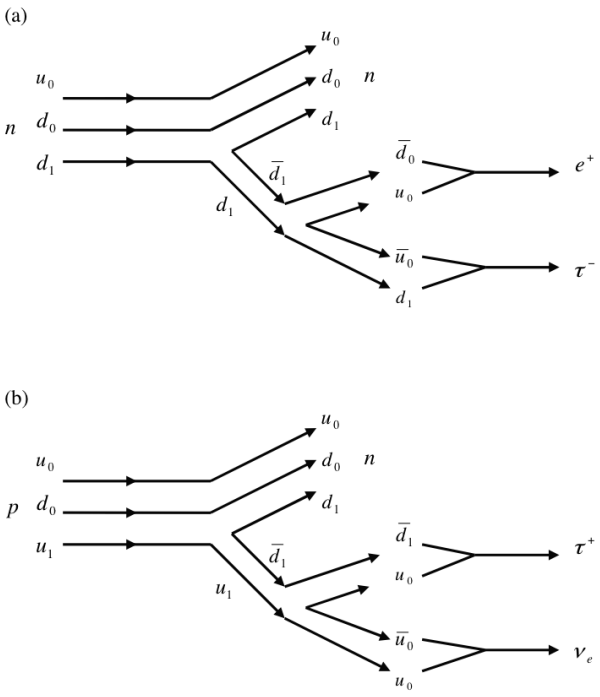


Fig. 6: Production of tau particles by a muon neutrino beam. (a) A negative tau particle is produced when an excited quark-antiquark pair, which is disintegrated from a muon neutrino and excited, interacts with a neutron. (b) A positive tau particle is produced when an excited quark-antiquark pair, which is disintegrated from a muon neutrino and excited, interacts with a proton.

cleon. Colliding with a neutron, a muon neutrino ν_μ is disintegrated into a ground state down-antidown quark pair ($d_0\bar{d}_0$), which can be excited into ($d_1\bar{d}_1$) and ($u_1\bar{u}_1$) when the flavor is also changed. The excited down-antidown quark pair ($d_1\bar{d}_1$) can either completely annihilate into a lambda neutrino ν_λ (Figure 5a) or interact with the neutron to generate a negative tau particle τ^- when the excited antidown quark degenerates into a ground state antidown and a ground state up-antiup quark pair, $\bar{d}_1 \longrightarrow \bar{d}_0 + (u_0\bar{u}_0)$ (Figure 6a). As shown in Figure 6a, the excited down quark in the neutron can incompletely annihilate with the ground state antiup quark into a negative tau particle τ^- and the ground state antidown quark can incompletely annihilate the ground state up quark into a positron e^+. Interacting with a proton (Figure 6b), the excited down-antidown quark pair ($d_1\bar{d}_1$) can generate a positive tau particle τ^+ when the excited up quark in the proton degenerates into a ground state up quark and a ground state up-antiup quark pair, $u_1 \longrightarrow u_0 + (u_0\bar{u}_0)$. The excited antidown quark can incompletely annihilate with the ground state up quark into a positive tau particle τ^+ and the ground state antiup quark can completely annihilate with the ground state up quark into an electron neutrino ν_e. If the excited up quark is not degenerated but directly annihilate with the excited antidown quark, a lambda particle λ^+ is produced (as shown in Figure 3).

On the other hand, for an electron neutrino beam, colliding with a nucleon, an electron neutrino ν_e is disintegrated into a ground state up-antiup quark pair ($u_0\bar{u}_0$) and excited into ($u_1\bar{u}_1$), which may be also from the disintegration of a muon neutrino with the favor change. This excited up-antiup quark-antiquark pair can either completely annihilate into a tau neutrino as shown in Figure 1 or interact with the nucleon to generate a muon particle (Figure 7). If the flavor is also changed, the annihilation and interaction with nucleons will produce the tau particles and neutrinos as shown in Figure 6 or lambda particles and neutrinos as shown in Figure 3.

Therefore, with the lepton formation and quark-antiquark pair production model developed in this paper, we can understand the recent measurement of a tau particle in a muon neutrino beam as well as the early measurements of muon particles in electron neutrino beams. More future experiments of the Large Electron-Positron Collider at CERN and measurements of neutrino oscillations are expected to validate this lepton formation and quark-antiquark pair production model and detect the fourth generation of leptons.

5 Conclusions

This paper develops a quark-antiquark annihilation and pair production model to explain the formation of leptons and the oscillation of neutrinos and further predict the fourth generation of leptons named as lambda particle and neutrino. It

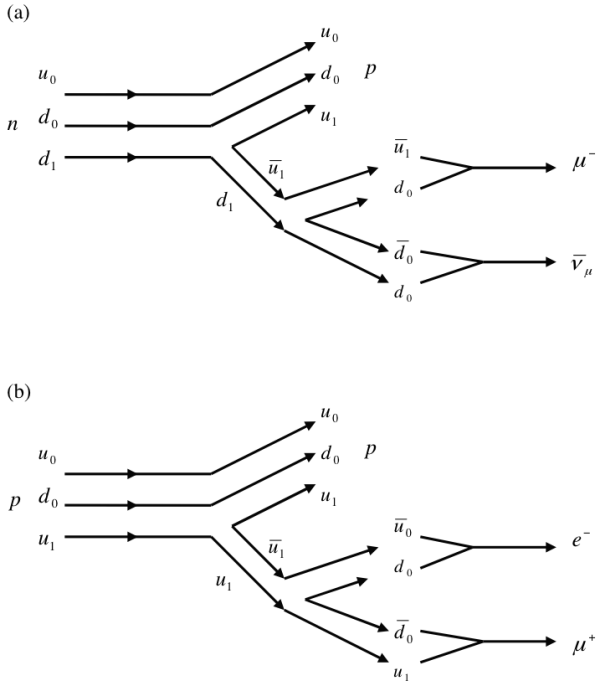


Fig. 7: Production of muon particles by an electron neutrino beam. (a) A negative muon particle is produced when an excited up-antiup quark pair, which is disintegrated from an electron neutrino and excited, interacts with a neutron. (b) A positive tau particle is produced when an excited quark-antiquark pair, which is disintegrated from a muon neutrino and excited, interacts with a proton.

is well known that all known or discovered leptons can be formed or emerged from particles or hadrons that are composed of only up and down quarks. This fact indicates that leptons must be consequences of activities of the up and down quarks and antiquarks. As quarks contain color charges, they participate in the strong interaction. Leptons do not contain color charges so that they do not participate in the strong interaction. In this paper, we suggested that all leptons are formed from quark-antiquark annihilations. There are two types of quark-antiquark annihilations. Type-I quark-antiquark annihilation annihilates only color charges, which forms structueless and colorless but electrically charged leptons such as electron, muon, tau, and lambda particles. Type-II quark-antiquark annihilation annihilates both electric and color charges, which forms structureless, colorless, and electrically neutral leptons such as electron, muon, tau, and lambda neutrinos. For the two types (up and down) of quarks and antiquarks to generate all four generation leptons from annihilations, they must have at least one excited state. Analyzing these two types of annihilations between up and down quarks and antiquarks with one excited quantum state for each of them, we predict the formation of the fourth generation of leptons named lambda particle and lambda neutrino. On the other hand, a lepton, when it collides with a nucleon, can be disintegrated into a quark-antiquark pair, which can be exited and/or changed in flavor. The quark-antiquark pair disintegrated from a neutrino can be excited and/or changed in flavor during the collision and then annihilate into another type of neutrino or interact with a nucleon to form electrically charged leptons. This quark-antiquark annihilation and pair production model provides a possible explanation for neutrino oscillations without hurting the standard model of particle physics. With it, we can understand the recent OPERA measurement of a tau particle in a muon neutrino beam as well as the early measurements of muon particles in electron neutrino beams.

Acknowledgement

This work was supported by the Title III program of Alabama A & M University, the NASA Alabama EPSCoR Seed grant (NNX07AL52A), and the National Natural Science Foundation of China (G40890161).

Submitted on January 20, 2011 / Accepted on January 26, 2011

References

1. Gell-Mann M. A schematic model of baryons and mesons. *Physics Letters*, 1964, v. 8, 214–215.
2. Zweig G. *CERN Report* No. 8182/TH.401, 1964.
3. Fritzsch H. Quarks. Basic Books, Inc., 1983.
4. Sher M., Yuan Y. Cosmological bounds on the lifetime of a fourth generation charged lepton. *Physics Letters B*, 1992, v. 285, 336–342.
5. Frampton P. H., Ng D., Sher M., Yuan Y. Search for heavy charged leptons at hadron colliders. *Physical Review D*, 1993, v. 48, 3128–3135.
6. Bhattacharyya G., Choudhury D. Fourth-generation leptons at LEP2. *Nuclear Physics B*, 1996, v. 468, 59-71.
7. Frandsen M. T., Masina I., Sannino F. Fourth lepton family is natural in technicolor. *Physical Review D*, 2010, v. 81, 35010.
8. Aguilar-Saavedra J. A. Heavy lepton pair production at LHC: Model discrimination with multi-lepton signals. *Nuclear Physics B*, 2010, v. 828, 289–316.
9. Nandi S., Rittenberg V., Schneider H. R. Quark-antiquark annihilation and small-pT inclusive spectra. *Physical Review D*, 1978, v. 17, 1336–1343.
10. Frenkel J., Gatheral J. G. M., Taylor J. C. Quark-antiquark annihilation is infrared safe at high energy to all orders. *Nuclear Physics B*, 1984, v. 233, 307–335.
11. Boros C., Liang Z. T., Meng T. C. Quark spin distribution and quark-antiquark annihilation in single-spin hadron-hadron collisions. *Physical Review Letters*, 1993, v. 70, 1751–1754.
12. Berrneuther W., Fucker M. and Si Z. G. Mixed QCD and Weak Corrections to tbar t Production by Quark-Antiquark Annihilation. *International Journal of Modern Physics A*, 2006, v. 21, 914–817.
13. Bredenstein A., Denner A., Dittmaier S., Pozzorini S. NLO QCD corrections to $t\bar{t}b\bar{b}$ production at the LHC: 1. quark-antiquark annihilation. *Journal of High Energy Physics*, 2008, v. 8, 108–108.
14. Zhang T. X. Electric charge as a form of imaginary energy. *Progress in Physics*, 2008, v. 2, 79–83.
15. Zhang T. X. Fundamental elements and interaction of the Nature: a classical unification theory. *Progress in Physics*, 2010, v. 2, 36–42.
16. DeGrand T. A., Jaffe R. L. Excited states of confined quarks. *Annals of Physics*, 1976, v. 100, 425–456.

17. de Rujula A., Maiani L., Petronzio R. Search for excited quarks. *Physics Letters B*, 1984, v. 140, 253–258.

18. Baur U., Spira M., Zerwas P. M. Excited-quark and -lepton production at hadron colliders. *Physical Review D*, 1990, v. 42, 815–824.

19. Wolfenstein L. Neutrino oscillations in matter. *Physical Review D*, 1978, v. 17, 2369–2374.

20. Aguilar A. et al. Evidence for neutrino oscillations from the observation of $\bar{\nu}_e$ appearance in a $\bar{\nu}_\mu$ beam. *Physical Review D*, 2001, v. 64, 112007.

21. Acquafredda R. et al. The OPERA experiment in the CERN to Gran Sasso neutrino beam. *Journal of Instrumentation*, 2009, v. 4, 4018.

22. Agafonova N. et al. Observation of a first ν candidate event in the OPERA experiment in the CNGS beam. *Physics Letters B*, 2010, v. 691, 139–145.

23. Dusini S. The OPERA experiment: A direct search of the $\nu_\mu \longrightarrow \nu_\tau$ oscillations. *Progress in Particle and Nuclear Physics*, 2010, v. 64, 187–189.

On the Independent Determination of the Ultimate Density of Physical Vacuum

Anatoly V. Belyakov

E-mail: belyakov.lih@gmail.com

In this paper, we attempt to present physical vacuum as a topologically non-unitary coherent surface. This representation follows with J. A. Wheeler's idea about fluctuating topology, and provides a possibility to express some parameters of the unit space element through the fundamental constants. As a result, we determined the ultimate density of physical vacuum without use of Hubble's constant.

The ultimate density of physical vacuum is regularly calculated through the experimentally obtained quantity — Hubble's constant. This constant follows from astronomical observations, and therefore its numerical value is under permanent update. On the other hand, the ultimate density of physical vacuum can also be determined in an independent way, through only the fundamental constants. Moreover, in the framework of this mechanistic model, it does not matter what we mean saying "physical vacuum": a material substance, or space itself.

As an initial model of the space micro-element of matter, it is reasonable to use J. A. Wheeler's idea about fluctuating topology. In particular, electric charges are considered therein as singular points located in a three-dimensional surface, and connected to each other through "wormholes" or current tubes of the input-output (source-drain) kind in an additional dimension, thus forming a closed contour. Is this additional dimension really required? It is probably that the three-dimensional space, if considered at a microscopic scale, has really lesser number of dimensions, but is topologically non-unitary coherent and consists of linkages [1].

The most close analogy to this model, in the scale of our world, could be the surface of an ideal liquid, vortical structures in it and their interactions which form both relief of the surface and sub-surface structures (vortical grids, etc.).

From the purely mechanistic point of view, this model should not contain the *electric charge* as a special kind of matter: the electric charge only manifests the degree of the non-equilibrium state of physical vacuum; it is proportional to the momentum of physical vacuum in its motion along the contour of the vortical current tube. Respectively, the *spin* is proportional to the angular momentum of the physical vacuum with respect to the longitudinal axis of the contour, while the *magnetic interaction* of the conductors is analogous to the forces acting among the current tubes.

Of course, in the framework of this model, both point and line are means physical objects, which have specific sizes. We assume the classical radius of the electron r_e as the minimal size. This approach was already justified in determination of the numerical value of the electron's charge, and the constants of radiation [2].

Thus matter itself can finally be organized with step-by-step complication of the initial contours, and be a "woven cloth", which, in its turn, is deformed into the objects we observe. The objects therefore are very fractalized (upto the micro-world scales) *surfaces*, which have a fractional dimension of the number almost approaching three and presupposingly equal to the number e.

The latter conclusion verifies that fact that the function $n^{1/n}$, which can be interpreted as the length of a ridge of the unit cube (its volume is equivalent to the summary volume of n cubes of the nth dimension), reach its maximum at exact $n = e$. We can understand this result so that the world of this dimension $n = e$ is most convex to the other worlds.

As a result, the surface being non-deformed can logically be interpreted as empty space, while the deformed and fractalized surface, i.e. the surface bearing an information about local deformations — as substance, masses. What is about an absolutely continuous three-dimensional body, it has not any internal structure thus does not bear any information about its interior (except, as probably, its own mass): such bodies do not really exist, even in the real micro world. In other words, the surface is material. However, being non-deformed, it does not manifest its material properties.

It should be noted that it is impossible to discuss the real geometry of topology of the world in the framework of this concept. Moreover, it is quite complicate to differ the surface from space, and space from matter, because such a step means at least a conceivable leaving our surface, which consists ourselves and the Universe itself. On the other hand, our model does not require such a step.

To calculate the density of physical vacuum in the framework of our suggested model, it is sufficient to determine the square, thickness, and mass of the "smoothed" surface (non-perturbed physical vacuum), then reduce the mass to the ultimate volume. To do it, we need to determine parameters of the initial micro-element and elementary contour.

According to the assumed model, we write down the electric forces F_e and the magnetic forces F_m in the "coulombless" form, where we replace the electric charge with the ultimate momentum of the electron $m_e c$. We obtain, for the electric forces,

$$F_e = \frac{z_1 z_2 (m_e c)^2}{\varepsilon_0 r^2}, \qquad (1)$$

where z_1 and z_2 are the numbers of the electric charges, c is the velocity of light, m_e is the electron's mass, while $\varepsilon = \frac{m_e}{r_e}$ is a new electric constant, which is 3.23×10^{-16} kg/m and is the linear density of the vortical tube. Respectively, for the magnetic forces, we obtain

$$F_m = \frac{z_1 z_2 \mu_0 (m_e c)^2 L}{2\pi r \times [\sec^2]}, \qquad (2)$$

where $\mu_0 = \frac{1}{\varepsilon_0 c^2}$ is a new magnetic constant, whose numerical value is 0.034 H^{-1}, L is the length of the conductors of the current (vortical tubes), while r is the distance between them. Numerically, the electric forces (1) and the magnetic forces (2) coincide with those calculated according to the standard equations of electrodynamics.

Thus, the quantity inverse to the magnetic constant, is the centrifugal force which appears due to the rotation of the vortical tube's element whose mass is m_e, with the velocity of light c around the radius r_e. The centrifugal force is also equivalent to the force acting among two elementary electric charges at this radius. We note, that the latter conclusion is the same as that W. C. Daywitt arrived at in the paper [3].

In the non-perturbed physical vacuum the electric, magnetic, and other forces should be compensated. In particular, proceeding from the equality of the electric and magnetic forces, one can deduce the *geometric mean value*, which is a linear parameter independent from the direction of the vortical tubes and the number of the electric charges

$$R_c = \sqrt{Lr} = \sqrt{2\pi}\, c \times [\sec] = 7.51 \times 10^8 \text{ [m]}. \qquad (3)$$

This fundamental length is close to the radius of the Sun and also the sizes of many typical stars.

Thus equation (3) represents the ratio between the contour's length and the distance between the vortical tubes. Now, assuming that the figures of the contours satisfy the formula (3), we are going to calculate the total mass of physical vacuum, which fills the Universe, and also its density.

Let the "smoothed" surface of physical vacuum has a size of $L \times L$, and is densely woven on the basis of parallel vortical tubes (they have parameters L and r) which, in their turn, are filled with the unit threads (each of the threads has a radius equal to the radius of the electron r_e). Also, assume that, even if there exist structures whose size is lesser than r_e, they do not change the longitudinal density ε_0. Thus, the total mass of the surface of the thickness r, which is the mass of physical vacuum M_v (including all hidden masses), is obviously determined by the formula

$$M_v = \frac{\pi}{4} \varepsilon_0 L \frac{L}{r} \left(\frac{r}{r_e}\right)^2. \qquad (4)$$

The average density of the substance of physical vacuum ρ_v is expressed through the ratio of the mass M_v to the spherical volume $\frac{4}{3}\pi L^3$. As a result, extending the formula of ε_0 then expressing L from (3), we obtain

$$M_v = \frac{\pi}{4} \frac{\rho_e R_c^4}{r}, \qquad (5)$$

$$\rho_v = \frac{3}{16} \rho_e \left(\frac{r}{R_c}\right)^2, \qquad (6)$$

where ρ_e is the density of the electron derived from its classical parameters, and is $\rho_e = \frac{m_e}{r_e^3} = 4.07 \times 10^{13}$ kg/m^3.

The main substance of the Universe is hydrogen. Therefore, it is naturally to assume r equal to the size of the standard proton-electron contour, which is the Bohr 1st radius 0.53×10^{-10} m.

Thus we obtain: the ultimate large length of the contour $L = 1.06 \times 10^{28}$ m, the total mass of substance in the Universe $M_v = 1.92 \times 10^{59}$ kg, and the ultimate density of physical vacuum $\rho_v = 3.77 \times 10^{-26}$ kg/m^3 (or $\rho_v = 3.77 \times 10^{-29}$ g/cm^3 in the CGS units).

The calculated numerical value of the average density of matter (physical vacuum) in the Universe is close to the modern esteems of the crucial density (the density of the *Einsteinian vacuum*).

With breaking the homogeneity of physical vacuum the anisotropy appears in the Universe. This is subjectively perceived in our world as manifestations of the pace of time, and the preferred directions in space. It is possible to suppose that the Universe as a whole evolutionary oscillates near its state of equilibrium, thus deforming the vacuum medium and creating the known forms of matter as a result. The stronger deformation, the larger contours (the heavier elements of substance) appear. Proceeding from the fact that elements with more than seven electron shells are unknown, we can conclude that the scale of the evolutionary oscillations of the Universe in the part of deformations of its own "tissue" is very limited. This is despite, as is probably, the Universe goes through the zero-state of equilibrium of physical vacuum during its evolution, where all real masses approach to zero, and the forces of gravitation — to their minimum. Here we see a relative connexion to *Mach's principle*, i.e. a dependency of the masses of objects on the mass of the entire Universe (of course, if meaning the mass of the entire Universe as the mass of physical vacuum, which is much greater than the summary mass of regular substance).

In conclusion, we suggest a supposition. Because masses or physical objects are merely forms of the relief of the surface of the vacuum medium, *which can only exist if the forming medium moves permanently along ordered trajectories*, in the framework of this interpretation *time* manifests evolution, change of objects and structures along the direction of motion of matter they consist of. Therefore, all paradoxes of time vanish here: the hypothetical displacement of an observer toward or backward with the current of matter leads only to his arrival at his alternative past or future; his actions

therein cannot change his own present — his own evolving section of the Universe.

Submitted on January 07, 2011 / Accepted on January 10, 2011

References

1. Dewitt B. S. Quantum gravity. *Scientific American*, v. 249, December 1983, 112–129.
2. Belyakov A. V. Charge of the electron, and the constants of radiation according to J. A. Wheeler's geometrodynamic model. *Progress in Physics*, 2010, v. 4, 90–94.
3. Daywitt W. C. The Relativity Principle: space and time and the Plank vacuum. *Progress in Physics*, 2010, v. 4, 34–35.

An Elegant Argument that $P \neq NP$

Craig Alan Feinstein

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

In this note, we present an elegant argument that $P \neq NP$ by demonstrating that the Meet-in-the-Middle algorithm must have the fastest running-time of all deterministic and exact algorithms which solve the SUBSET-SUM problem on a classical computer.

Consider the following problem: Let $A = \{a_1, \ldots, a_n\}$ be a set of n integers and b be another integer. We want to find a subset of A for which the sum of its elements (we shall call this quantity a *subset-sum*) is equal to b (we shall call this quantity the *target integer*). We shall consider the sum of the elements of the empty set to be zero. This problem is called the *SUBSET-SUM problem* [1,2]. Now consider the following algorithm for solving the SUBSET-SUM problem:

Meet-in-the-Middle Algorithm - First, partition the set A into two subsets, $A^+ = \{a_1, \ldots, a_{\lceil \frac{n}{2} \rceil}\}$ and $A^- = \{a_{\lceil \frac{n}{2} \rceil + 1}, \ldots, a_n\}$. Let us define S^+ and S^- as the sets of subset-sums of A^+ and A^-, respectively. Sort sets S^+ and $b - S^-$ in ascending order. Compare the first elements in both of the lists. If they match, then output the corresponding solution and stop. If not, then compare the greater element with the next element in the other list. Continue this process until there is a match, in which case there is a solution, or until one of the lists runs out of elements, in which case there is no solution.

This algorithm takes $\Theta(\sqrt{2^n})$ time, since it takes $\Theta(\sqrt{2^n})$ steps to sort sets S^+ and $b - S^-$ and $O(\sqrt{2^n})$ steps to compare elements from the sorted lists S^+ and $b - S^-$. It turns out that no deterministic and exact algorithm with a better worst-case running-time has ever been found since Horowitz and Sahni discovered this algorithm in 1974 [3, 4]. In this paper, we shall prove that it is impossible for such an algorithm to exist.

First of all, we shall assume, without loss of generality, that the code of any algorithm considered in our proof does not contain full or partial solutions to any instances of SUBSET-SUM. This is because only finitely many such solutions could be written in the code, so such a strategy would not be helpful in speeding up the running-time of an algorithm solving SUBSET-SUM when n is large. We now give a definition:

Definition 1: *We define a γ-comparison (a generalized comparison) between two integers, x and y, as any finite procedure that directly or indirectly determines whether or not $x = y$.*

For example, a finite procedure that directly compares $f(x)$ and $f(y)$, where x and y are integers and f is a one-to-one function, γ-compares the two integers x and y, since $x = y$ if and only if $f(x) = f(y)$. Using this expanded definition of compare, it is not difficult to see that the Meet-in-the-Middle algorithm γ-compares subset-sums with the target integer until it finds a subset-sum that matches the target integer. We shall now prove two lemmas:

Lemma 2: *Let x and y be integers. If $x = y$, then the only type of finite procedure that is guaranteed to determine that $x = y$ is a γ-comparison between x and y. And if $x \neq y$, then the only type of finite procedure that is guaranteed to determine that $x \neq y$ is a γ-comparison between x and y.*

Proof: Suppose there is a finite procedure that is guaranteed to determine that $x = y$, if $x = y$. Then if the procedure does not determine that $x = y$, this implies that $x \neq y$. Hence, the procedure always directly or indirectly determines whether or not $x = y$, so the procedure is a γ-comparison between x and y.

And suppose there is a finite procedure that is guaranteed to determine that $x \neq y$, if $x \neq y$. Then if the procedure does not determine that $x \neq y$, this implies that $x = y$. Hence, the procedure always directly or indirectly determines whether or not $x = y$, so the procedure is a γ-comparison between x and y. ■

Lemma 3: *Any deterministic and exact algorithm that finds a solution to SUBSET-SUM whenever a solution exists must (whenever a solution exists) γ-compare the subset-sum of the solution that it outputs with the target integer.*

Proof: Let Q be a deterministic and exact algorithm that finds a solution, $\{a_{k_1}, \ldots, a_{k_m}\}$, to SUBSET-SUM whenever a solution exists. Before Q outputs this solution, it must verify that $a_{k_1} + \ldots + a_{k_m} = b$, since we are assuming that the code of Q does not contain full or partial solutions to any instances of SUBSET-SUM. (See above for an explanation.) In order for Q to verify that $a_{k_1} + \ldots + a_{k_m} = b$, Q must γ-compare the subset-sum, $a_{k_1} + \ldots + a_{k_m}$, with the target integer, b, since a γ-comparison between $a_{k_1} + \ldots + a_{k_m}$ and b is the only type of finite procedure that is guaranteed to determine that $a_{k_1} + \ldots + a_{k_m} = b$, by Lemma 2. ■

As we see above, the Meet-in-the-Middle algorithm makes use of sorted lists of subset-sums in order to obtain a faster

worst-case running-time, $\Theta(\sqrt{2^n})$, than that of a naïve brute-force search of the set of all subset-sums, $\Theta(2^n)$. The following lemma shows that sorted lists of subset-sums are necessary to achieve such an improved worst-case running-time.

Lemma 4: *If a deterministic and exact algorithm that finds a solution to SUBSET-SUM whenever a solution exists does not make use of sorted lists of subset-sums, it must run in* $\Omega(2^n)$ *time in the worst-case scenario.*

Proof: Let Q be a deterministic and exact algorithm that finds a solution to SUBSET-SUM whenever a solution exists without making use of sorted lists of subset-sums. By Lemma 3, Q must (whenever a solution exists) γ-compare the subset-sum of the solution that it outputs with the target integer. In order for Q to avoid wasting time γ-comparing the target integer with subset-sums that do not match the target integer, Q must be able to rule out possible matches between subset-sums and the target integer without γ-comparing these subset-sums with the target integer.

But by Lemma 2, the only type of finite procedure that is guaranteed to rule out a possible match between a subset-sum and the target integer, if they do not match, is a γ-comparison. So in the worst-case scenario, there is no way for Q to avoid wasting time γ-comparing the target integer with subset-sums that do not match the target integer. And since Q does not make use of sorted lists of subset-sums like the Meet-in-the-Middle algorithm does, its γ-comparisons between these subset-sums and the target integer will not rule out possible matches between any other subset-sums and the target integer. Hence, in the worst-case scenario Q must γ-compare each of the 2^n subset-sums with the target integer, which takes $\Omega(2^n)$ time. ■

It is now possible, using Lemma 4, to solve the problem of finding a nontrivial lower-bound for the worst-case running-time of a deterministic and exact algorithm that solves the SUBSET-SUM problem:

Theorem 5: *It is impossible for a deterministic and exact algorithm that solves the SUBSET-SUM problem to have a worst-case running-time of* $o(\sqrt{2^n})$.

Proof: Let T be the worst-case running-time of an algorithm Q that solves SUBSET-SUM, and let M be the size of the largest list of subset-sums that Q sorts. Since it is necessary for Q to make use of sorted lists of subset-sums in order to have a faster worst-case running-time than $\Theta(2^n)$ by Lemma 4 and since it is possible for Q to make use of sorted lists of size M of subset-sums to rule out at most M possible matches of subset-sums with the target integer at a time (just as the Meet-in-the-Middle algorithm does, with $M = \Theta(\sqrt{2^n})$), we have $MT \geq \Theta(2^n)$. And since creating a sorted list of size M must take at least M units of time, we have $T \geq M \geq 1$.

Then in order to find a nontrivial lower-bound for the worst-case running-time of an algorithm solving SUBSET-SUM, let us minimize T subject to $MT \geq \Theta(2^n)$ and $T \geq M \geq 1$. Because the running-time of $T = \Theta(\sqrt{2^n})$ is the solution to this optimization problem and because the Meet-in-the-Middle algorithm achieves this running-time, 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 $P \neq NP$ [1, 5]. ■

Acknowledgments

I thank God, my parents, my wife, and my children for their support.

Submitted on January 11, 2011 / Accepted on January 15, 2011

References

1. Cormen T. H., Leiserson C. E., Rivest R. L. Introduction to Algorithms, McGraw-Hill, 1990.
2. Menezes A., van Oorschot P., Vanstone S. Handbook of Applied Cryptography, CRC Press, 1996.
3. Horowitz E., Sahni S. Computing Partitions with Applications to the Knapsack Problem. *Journal of the ACM*, v. 2l, no. 2, April 1974, pp. 277–292.
4. Woeginger G. J. Exact Algorithms for NP-Hard Problems, *Lecture Notes in Computer Science*, 2003, v. 2570, pp. 185–207.
5. Bovet P. B., Crescenzi P. Introduction to the Theory of Complexity, Prentice Hall, 1994.

The Compton Radius, the de Broglie Radius, the Planck Constant, and the Bohr Orbits

William C. Daywitt

National Institute for Standards and Technology (retired), Boulder, Colorado, USA
E-mail: wcdaywitt@earthlink.net

The Bohr orbits of the hydrogen atom and the Planck constant can be derived classically from the Maxwell equations and the assumption that there is a variation in the electron's velocity about its average value [1]. The resonant nature of the circulating electron and its induced magnetic and Faraday fields prevents a radiative collapse of the electron into the nuclear proton. The derived Planck constant is $h = 2\pi e^2/\alpha c$, where e, α, and c are the electronic charge, the fine structure constant, and the speed of light. The fact that the Planck vacuum (PV) theory [2] derives the same Planck constant independently of the above implies that the two derivations are related. The following highlights that connection.

In the Beckmann derivation [1], the electromagnetic-field mass and the Newtonian mass are assumed to have the same magnitude in which case the electron's average kinetic energy can be expressed as

$$\left(\frac{mv^2}{2}\right)_{\text{em}} + \left(\frac{mv^2}{2}\right)_{\text{n}} = mv^2 = mv\cdot v = mv\cdot\lambda\nu = mv\lambda\cdot\nu = h\nu \quad (1)$$

where v is the average electron velocity and $v = \lambda\nu$ is a simple kinematic relation expressing the fact that the electron's instantaneous velocity varies periodically at a frequency ν over a path length equal to the wavelength λ. The constant $h\,(= mv\lambda)$ turns out to be the Planck constant.

The Beckmann derivation assumes with Maxwell and those following thereafter that the magnetic and Faraday fields are part of the electron makeup. On the other hand the PV theory assumes that these fields constitute a reaction of the negative-energy PV quasi-continuum to the movement of the massive point charge (the Dirac electron). In its rest frame the electron exerts the two-fold force [3]

$$\frac{e_*^2}{r^2} - \frac{mc^2}{r} \quad (2)$$

on each point r of the PV, where e_* $(= e/\sqrt{\alpha})$ is the electron's bare charge, e is the laboratory-observed charge, and m is the electron mass. The vanishing of this composite force at the radius $r = r_c$ leads to

$$r_c mc^2 = e_*^2 = c\hbar = e^2/\alpha, \quad (3)$$

where r_c is the electron's Compton radius and $\hbar$ is the (reduced) Planck constant. From the introductory paragraph and (3), the Beckmann and PV results clearly lead to the same Planck constant $\hbar = e^2/\alpha c = e_*^2/c$.

The Planck constant then is associated only with the bare charge $|e_*|$ and not the electron mass—thus the quantum theory reflects the fact that, although the various elementary particles have different masses, they are associated with only one electric charge.

The expression $mv\lambda = h$ used in (1) to arrive at the total electron kinetic energy is the de Broglie relation expressed in simple, physically intuitive terms: the de Broglie relation yields the product of the electron mass m, its average velocity v, and the path length λ over which its instantaneous velocity varies. The relativistic version of the relationship (which is arrived at in the Appendix by assuming the vanishing of (2) at $r = r_c$ to be a Lorentz invariant constant) is

$$m\gamma v = \frac{\hbar}{r_d} = \frac{\gamma\hbar}{r_c/\beta} = \frac{\gamma h}{\lambda_c/\beta} = \frac{\gamma h}{\lambda} \quad (4)$$

where $m\gamma v$ is the relativistic momentum; and $\lambda = \lambda_c/\beta$, where λ_c is the Compton wavelength $2\pi r_c$. Thus Beckmann's de Broglie relation is in relativistic agreement with the PV result.

The preceding demonstrates that Bohr's introduction of the quantum concept in terms of an ad-hoc Planck constant [4] can be derived from classical electromagnetism and the assumption that the electron interacts with some type of negative-energy vacuum state (the PV in the present case). That the Lorentz transformation can also be derived from the same assumptions is shown in a previous paper [5].

Acknowledgment

The present author's first contact with the late Professor Petr Beckmann was in a course he taught at the University of Colorado (USA) around 1960 on 'Statistical Communication Theory' and later (~circa 1989) in a number of phone conversations concerning his book *Einstein Plus Two* [1]. Much of the work on the PV theory was inspired by Prof. Beckmann's relentless search for the physical truth of things. In addition to authoring a number of interesting books, he founded the scientific journal *Galilean Electrodynamics* and the news letter *Access to Energy* both of which are still active today.

Appendix: de Broglie Radius

The Dirac electron exerts two distortion forces on the collection of Planck particles constituting the degenerate PV, the

polarization force e_*^2/r^2 and the curvature force mc^2/r. The equality of the two forces at the electron Compton radius r_c is assumed to be a fundamental property of the electron-PV interaction. The vanishing of the force difference $e_*^2/r_c^2 - mc^2/r_c = 0$ (a Lorentz invariant constant) at the Compton radius can be expressed as a vanishing 4-force difference tensor [6]. In the primed rest frame of the electron, where these static forces apply, this force difference $\Delta F'_\mu$ is

$$\Delta F'_\mu = \left[\mathbf{0}, i\left(\frac{e_*^2}{r_c^2} - \frac{mc^2}{r_c}\right)\right] = [0, 0, 0, i\,0] \qquad \text{(A1)}$$

where $i = \sqrt{-1}$. Thus the vanishing of the 4-force component $\Delta F'_4 = 0$ in (A1) is the Compton-radius result from (2) and can be expressed in the form $mc^2 = e_*^2/r_c = (e_*^2/c)(c/r_c) = \hbar\omega_c$, where $\omega_c \equiv c/r_c = mc^2/\hbar$ is the corresponding Compton frequency.

The 4-force difference in the laboratory frame, $\Delta F_\mu = a_{\mu\nu}\Delta F'_\nu = 0_\mu$, follows from its tensor nature and the Lorentz transformation $x_\mu = a_{\mu\nu}\, x'_\nu$ [6], where $x_\mu = (x, y, z, ict)$,

$$a_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -i\beta\gamma \\ 0 & 0 & i\beta\gamma & \gamma \end{pmatrix} \qquad \text{(A2)}$$

$\gamma = 1/\sqrt{1-\beta^2}$, and $\mu, \nu = 1, 2, 3, 4$. Thus (A1) becomes

$$\Delta F_\mu = \left[0, 0, \beta\gamma\left(\frac{e_*^2}{r_c^2} - \frac{mc^2}{r_c}\right), i\,\gamma\left(\frac{e_*^2}{r_c^2} - \frac{mc^2}{r_c}\right)\right]$$

$$= \left[0, 0, \left(\frac{e_*^2}{\beta\gamma r_d^2} - \frac{mc^2}{r_d}\right), i\left(\frac{e_*^2}{\gamma r_L^2} - \frac{mc^2}{r_L}\right)\right] = [0, 0, 0, i\,0] \qquad \text{(A3)}$$

in the laboratory frame. The equation $\Delta F_3 = 0$ from the final two brackets yields the de Broglie relation

$$p = \frac{e_*^2/c}{r_d} = \frac{\hbar}{r_d} \qquad \text{(A4)}$$

where $p = m\gamma v$ is the relativistic electron momentum and $r_d \equiv r_c/\beta\gamma$ is the de Broglie radius.

The equation $\Delta F_4 = 0$ from (A3) leads to the relation $p = \hbar/r_L$, where $r_L \equiv r_c/\gamma$ is the length-contracted r_c in the *ict* direction. The Synge primitive quantization of flat spacetime [7] is equivalent to the force-difference transformation in (A3): the ray trajectory of the particle in spacetime is divided (quantized) into equal lengths of magnitude $\lambda_c = 2\pi r_c$ (this projects back on the '*ict*' axis as $\lambda_L = 2\pi r_L$); and the de Broglie wavelength calculated from the corresponding spacetime geometry. Thus the development in the previous paragraphs provides a physical explanation for Synge's spacetime quantization in terms of the two perturbations e_*^2/r^2 and mc^2/r the Dirac electron exerts on the PV.

Submitted on February 01, 2011 / Accepted on February 05, 2011

References

1. Beckmann P. Einstein Plus Two, The Golem Press, Boulder, Colorado, 1987, (see Chapter 2).
2. Daywitt W.C. The Planck vacuum, *Progress in Physics*, 2009, v. 1, 20–26.
3. Daywitt W.C. The Dirac Electron in the Planck Vacuum Theory, *Progress in Physics*, 2010, v. 4, 69–71.
4. Leighton R. B. Principles of Modern Physics, McGraw-Hill Book Co., New York, 1959.
5. Daywitt W. C. The Lorentz Transformation as a Planck Vacuum Phenomenon in a Galilean Coordinate System, *Progress in Physics*, 2011, v. 1, 3–6.
6. Jackson J.D. Classical Electrodynamics, John Wiley & Sons, 1st ed., 2nd printing, New York, 1962.
7. Synge J.L. Geometrical Mechanics and de Broglie Waves, Cambridge University Press, 1954, (see pages 106–107).

Histograms Constructed from the Data of ^{239}Pu Alpha-Activity Manifest a Tendency for Change in the Similar Way as at the Moments when the Sun, the Moon, Venus, Mars and Mercury Intersect the Celestial Equator

Simon E. Shnoll[*†§], Ilya A. Rubinstein[‡], Sergey N. Shapovalov[♯], Valery A. Kolombet[†], Dmitri P. Kharakoz[†]

[*]Department of Physics, Moscow State University, Moscow, Russia
[†]Institute of Theor. and Experim. Biophysics, Russ. Acad. Sci., Pushchino, Russia
[‡]Skobeltsin Inst. of Nuclear Physics at Moscow State Univ., Moscow, Russia
[§]Pushchino State University, Pushchino, Russia
[♯]Arctic and Antarctic Research Institute, St. Petersburg, Russia

E-mail: shnoll@mail.ru (Simon E. Shnoll, the corresponding author)

Earlier, the shape of histograms of the results of measurements obtained in processes of different physical nature had been shown to be determined by cosmophysical factors [1]. Appearance of histograms of a similar shape is repeated periodically: these are the near-a-day, near-27-days and annual periods of increased probability of the similar shapes. There are two distinctly distinguished near-a-day periods: the sidereal-day (1,436 minutes) and solar-day (1,440 minutes) ones. The annual periods are represented by three sub-periods: the "calendar" (365 average solar days), "tropical" (365 days 5 hours and 48 minutes) and "sidereal" (365 days 6 hours and 9 minutes) ones. The tropical year period indicates that fact that histogram shape depends on the time elapsed since the spring equinox [2]. The latter dependence is studied in more details in this work. We demonstrate that the appearance of similar histograms is highly probable at the same time count off from the moments of equinoxes, independent from the geographic location where the measurements had been performed: in Pushchino, Moscow Region (54° NL, 37° EL), and in Novolazarevskaya, Antarctic (70° SL, 11° EL). The sequence of the changed histogram shapes observed at the spring equinoxes was found to be opposite to that observed at the autumnal equinoxes. As the moments of equinoxes are defined by the cross of the celestial equator by Sun, we also studied that weather is not the same as observed at the moments when the celestial equator was crossed by other celestial bodies — the Moon, Venus, Mars and Mercury. Let us, for simplicity, refer to these moments as a similar term "planetary equinoxes". The regularities observed at these "planetary equinoxes" had been found to be the same as in the case of true solar equinoxes. In this article, we confine ourselves to considering the phenomenological observations only; their theoretical interpretation is supposed to be subject of further studies.

1 Materials and methods

A many-year monitoring of the alpha-activity of ^{239}Pu samples (performed with devices constructed by one of the authors, I. A. Rubinstein [3]) was used as a basis for this study. Round-the-clock once-a-second measurements were made in Pushchino, at the Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences, and at Novolazarevskaya Antarctic Polar Station of the Arctic and Antarctic Institute.

Semi-conductor detectors used were either collimator-free or equipped with collimators limiting the beam of registered alpha-particles by a spatial angle (about 0.1 radian) within which the particles travelled along a certain direction: towards the Sun, Polar Star, West or East. The number of registered alpha-particles during one-second interval was the measured parameter. Results of the continuous once-a-second measurements of the decay activity were stored in a computer databank.

One-minute histograms constructed from 60 results of the once-a-second measurements of activity were the main objects of analysis in our study. The histograms were visually compared with each other in order to estimate the resemblance of their shape. The estimation was made by the method of expert judgment. This analysis was performed with the assistance of Edwin Pozharsky's computer program (described in [1]) which allowed to construct histograms for each one-minute interval in a series of measurements and, further, to smooth and scale them, and to mirror (if needed) in order to superimpose the histograms and visually compare their shape. At the final step of analysis, we constructed the distribution of the number of pairs of similar histograms versus the interval separating the histograms in each pair. Fig. 1 presents a diagram explaining three kinds of comparison of the series of histograms.

Method A; direct alignment (parallel). These two compared histogram series are aligned with each other as parallel

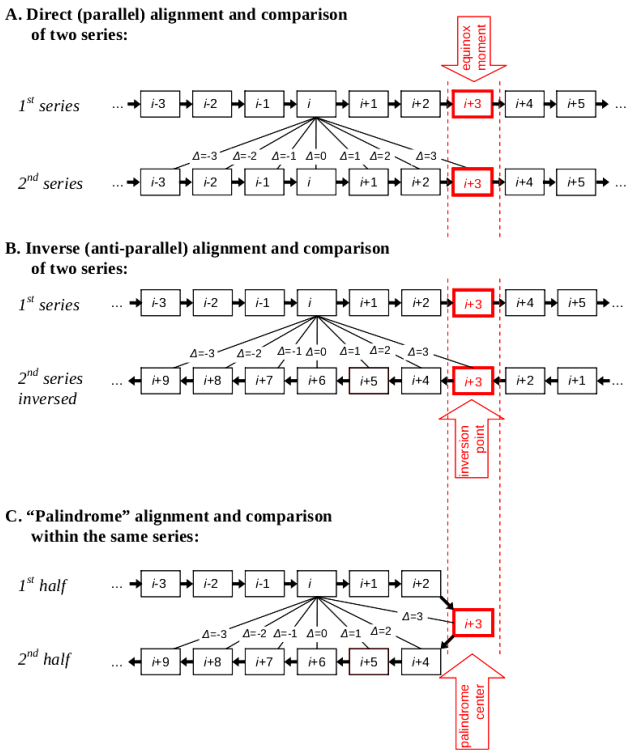


Fig. 1: Boxes indicate histograms constructed for the measured radioactivity within each minute of time. Arrows connecting the boxes indicate the direction of time separating the one-minute histograms. The interval between the histograms, Δ, is measured in minutes or in position numbers in the aligned series.

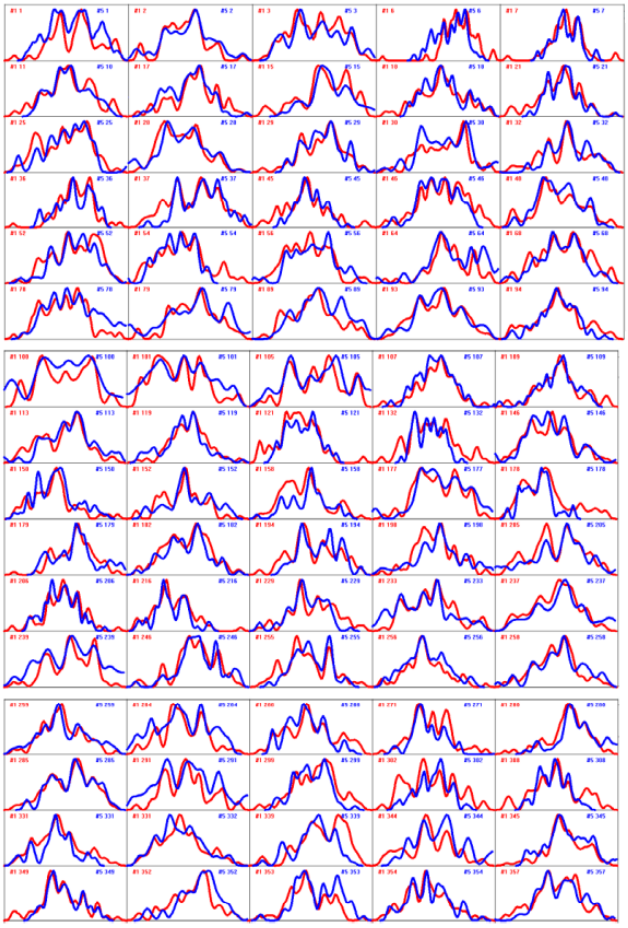


Fig. 2: Extract from a laboratory log: pairs of similar histograms of ^{239}Pu alpha-activity taken at the same time count off from the moments of Moon-2005-March-11 and Venus-2001-October-18 equinoxes. Each histogram is constructed by 60 one-second measurements (total time being equal to 1 minute). The doubled figures in upper corners of each picture specify the series number (No.1 or No.5) and the position number of a histogram in the series. Histograms of Moon-2005-March-11 equinox are drawn in red and those of Venus-2001-October-18 in blue. The data presented here correspond to the extreme in the blue curve in Fig. 4.

sequences in which the equinox moments occur at the same place. Then each i-th histogram of one series is compared with the number of neighboring histograms of the other series, as shown in Fig. 1, case A.

Method B; inverse alignment (anti-parallel). These two series are aligned in the same way but the second one is reversed at the point of equinox. This is illustrated in Fig. 1B.

Method C; "palindrome" alignment. Two parts of the same sequence are compared with each other. To do that, we assume the equinox moment to be the inversion center of a palindrome. Therefore, the second half of the sequence (following the center) is reversed and aligned with the first half as shown in Fig. 1C. Then the two halves of the sequence are compared with each other.

Fig. 2 presents an extract from the laboratory log-file to illustrate what shapes are considered similar from the expert's viewpoint. Final results are presented as the plots of the frequency of similar pairs of histograms versus the interval (measured in minutes) separating the position of items in the pairs (Fig. 3, e.g.).

The true equinox moments and the equivalent moments when the Moon, Mars, Venus or Mercury intersect the celestial equator (called here, by analogy, "planetary equinoxes") were determined by nonlinear interpolation of the data tabulated in the annual astronomy tables [4]; the residual error of this interpolation was much within the time resolution of our observations.

2 Results

2.1 Comparison of the histogram series obtained at the successive (in turn) "equinoxes"

It had been found earlier [2, 5, 6], in studying the variation of shape of the successive sequences of histograms obtained at vernal and autumnal equinoxes, the sequences related to "homonymous" equinoxes (vernal-and-vernal or autumnal-and-autumnal ones) display the similarity higher than that of "heteronymous" equinoxes. This fact gave a hint for a hypothesis that the histogram shape may depend not only on the

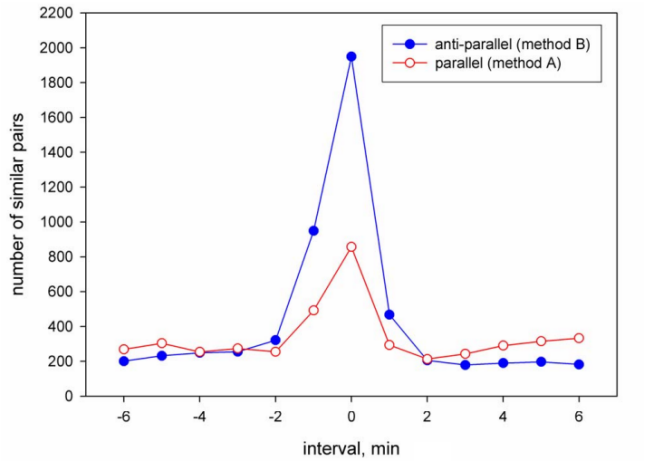


Fig. 3: Total number of similar pairs of histogram as function of time interval separating the histograms in the pair. The compared histogram sequences were related to the successive (neighboring) equinoxes. The sequences compared were aligned parallel or anti-parallel around the very "equinox" moment. First what is seen in the figure is that the probability to find a similar pair of histograms sharply decreases with the distance between the items in the pair. Second important moment is that the maximum probability (which is a measure of similarity of the sequences) is three times higher in the case the sequences compared are oriented anti-parallel. That means the histogram shape depends not only on the proximity of a planet or Sun to the celestial equator but also on the direction a celestial body moves to it. The diagram represents a summary of data for 24 pairs of compared series of alpha-activity aligned around the Sun, the Moon, Mars, Venus and Mercury "equinoxes". The series were 720 minutes long each. The data used have been obtained either in Pushchino and Novolazarevskaya.

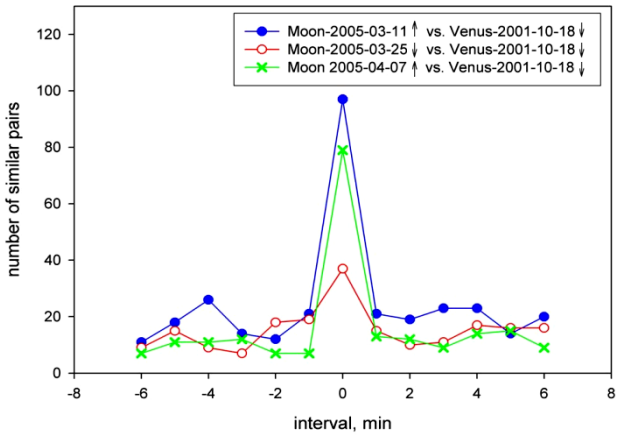


Fig. 4: Histogram series at the Venus setting "equinox" 2001-October-18(↓) is highly similar to those at the Moon rising "equinoxes" 2005-March-11(↑), 2005-April-7(↑) and considerably less similar to the series at the Moon setting "equinox" 2005-March-25(↓). (The extreme of this distribution is formed with the pairs of histograms presented in Fig. 2, to illustrate the extent of their similarity.)

proximity to the equinox moment, but also on the direction the Sun is moving to the celestial equator: from the northern or from the southern hemisphere. The hypothesis has been confirmed by comparison of the *direct* series of the autumn histograms with the inverse series of the spring histograms [2, 5, 6] (cf. in Fig. 1, method B).

In this work, the same analysis is applied to the data including "planetary equinoxes". Namely, not only solar but also lunar, venusian, martian and mercurian "equinoxes" have been considered. Fig. 3 summarizes the results of 24 pairs of such "equinox" events. Only the pairs of *successive* "equinoxes" were here compared; i.e., a series of histograms obtained at one "equinox" was compared with that of just next "equinox" for the same celestial body. The series were compared with use of either procedure A (parallel) and B (anti-parallel) (cf. in Fig. 1).

Fig. 3 shows that:

1. The phenomenon of similarity of the temporal change of the histogram shape near the moments of intersection of the celestial equator by a celestial body appears to be independent of the nature of the body (whatever the Sun, the Moon, Venus, Mars or Mercury). Hence, the sequential changes of histograms do not show any gravitational influence of the bodies;
2. Neither does it depend on the geographic point where the measurements have been performed: the data obtained in Pushchino and Novolazarevskaya (Antarctic) display similarity of histograms at the same absolute time, to within one-minute accuracy;
3. Neither does it depend on the velocity a celestial body moves across the equator. Despite a great difference in both angular and linear velocities of the bodies, the changes of histogram shape are correlated in the sequences aligned along the same time scale. This is a surprising result that has not a simple kinematic explanation. This fact, again, is an indication of that the phenomenon observed is not a matter of any "influence" exerted by a celestial bodies on the observable value;
4. What we can learn from the fact that the similarity of histogram sequences is higher in the case of anti-parallel orientation of compared sequences is as follows. The variation of histogram shape shows not only the extent of proximity of the Sun or a planet to the celestial equator but also their location in one or another celestial hemisphere. In other words, the "northern" histograms of a vernal equinox display higher similarity to the "southern" histograms of the next autumnal equinox despite the movement of the body on the sky sphere are reciprocal at these two cases.

The fact that the histogram shape variations do not depend on the nature of celestial body is confirmed also by pairwise comparison of the "equinoxes" of different planets. For example, Fig. 4 presents the results of comparison of a Venus "equinox" (2001-October-18 (↓)) with three Moon "equi-

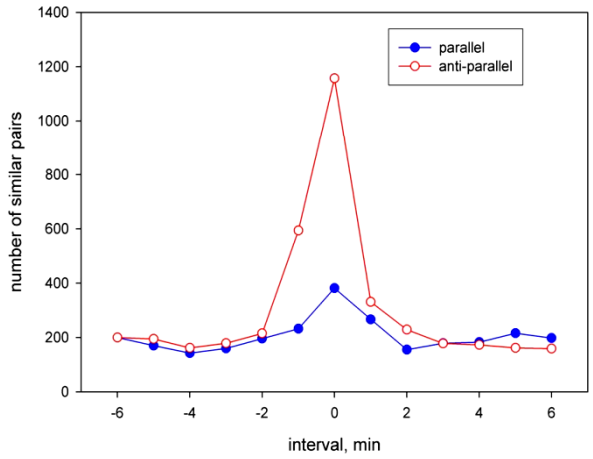


Fig. 5: The palindrome effect. The number of similar pairs of histograms in two compared halves of a sequence separated by the equinox moment as function of time shift between the halves. Total numbers obtained in the analysis of 17 equinox and "planetary equinox" events are presented here. Comparison of half-sequences was performed with the procedure C in Fig. 1. Palindrome effect is indicated by the fact that the similarity of anti-parallel half-sequences is 3-fold higher then that of parallel half-sequences.

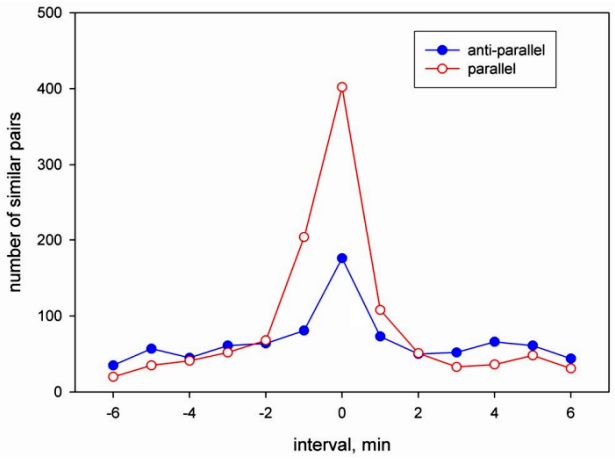


Fig. 6: Summarized data on 8 different Moon, Sun and Mars "equinoxes" when the ^{239}Pu alpha-activity was measured with the collimators directed towards Polar Star. Parallel oriented sequences of histograms display 3-fold higher similarity than those oriented anti-parallel.

noxes" (2005-March-11(↑); 2005-March-25(↓), and 2005-April-7(↑)). The figure shows that:

1. The histogram series adjusted to "equinox" moments are similar even if the events considered are separated by years (four years in this particular case);
2. The similarity of histograms depends on the direction of movement towards the celestial equator. However, in this particular case, symbate "equinoxes" (Venus-2001-October-18(↓) and Moon-2005-March-25(↓)) happened to be less similar then the counter-directed ones (Venus-2001-October-18(↓) and the other two Moon "equinoxes", both rising).

2.2 Comparison of direct and inverse halves of the same series of histograms (the "palindrome" effect)

The "palindrome effect" has been described in a number of earlier works; this is the presence of specific inversion points in the time series of histograms after which the same histograms occur in the reverse order [2, 5, 6]. Over a daily period, 6 am and 6 pm of local time have been found to be such inversion points. Any point on the Earth's surface participates in two movements, one due to the rotation of the Earth about its axis and another due to the movement of the center of the Earth along its circumsolar orbit. One finds that, the projections of the two movements onto the circumsolar orbit are counter-directional during the daytime (6 am to 6 pm) and co-directional during the nighttime (6 pm to 6 am). This is a probable "kinematic" reason for these two time moments are featured.

A question arose if there are a number of such "palindrome" centers existing at the equinoxes and "planetary equinoxes", separating the celestial body movement towards and away from the celestial equator. This supposition has proved to be true. Indeed, the histogram sequences preceding the equinox moments were similar to the inverse sequences following this moments. The palindrome effect (detected with the procedure C; cf. Fig. 1), is illustrated in Fig. 5. As revealed from the picture, the probability to find similar histograms is 3-fold higher if one of the compared half-sequences is inverted — this is exact the essence of the palindrome effect.

Phenomenological meaning of this observation is that the "equinox" moments are just the points at which the order of changing the histogram shapes is reverted. Sharpness of the peak is an indication of a high accuracy (one minute) of detection of the "equinox" moment. Mirror symmetry of the histogram sequences around an equinox means that histogram shape depends on how far distant is the celestial body from the equator plane, whatever on the northern or southern side of the celestial sphere. This result might seem to conflict with the data presented in Fig. 3 indicating that the direction of the movement also does matter. However, the seeming conflict is resolved by the fact (revealed from our special investigation) that these two phenomena are determined by different, not overlapping sets of similar histograms.

2.3 Equinox effects in the data obtained with collimators

The phenomena presented in Figs. 3-5 have been observed in the data on the alpha-activity measured with either collimator-free detectors and those supplied with the collimator permanently directed towards West or East or towards the Sun. No dependence of the observed phenomenology on the orientation of the collimator was found in these cases. The only

dramatic difference has been found for the case where the collimator was directed towards the Polar Star.

In this series of data, the sequences of histograms obtained at the sequential (neighboring) "equinoxes" displayed higher similarity when they were compared as parallel sequences (procedure A in Fig. 1) and lesser similarity when they were anti-parallel. This observation is illustrated in Fig. 6, summarizing the results of eight equinox events with the Moon, the Sun and Mars. Therefore, the difference between the northern and southern hemispheres is not reflected in the measurements with the collimator directed to Northern Pole.

3 Discussion

A number of phenomenological conclusions follow from the results presented herein:

1. The shape of histograms obtained from the measurements performed in different geographic locations near the time of "equinoxes" is changing synchronously, within one minute accuracy. For instance, they occur simultaneously in Pushchino and Novolazarevskaya despite 104 minutes of local time difference between these two places. This means this is a global phenomenon independent of the Earth axial rotation;
2. The changes of histograms obtained near the solar or planetary "equinoxes" do not depend on the nature of the "acting" celestial body, whatever the Sun, Mercury, Mars, Venus or Mercury. Individual features of the bodies — different masses and different rates of their orbital movement — are not essential;
3. At very equinox moments, the inversion of the sequence of the histogram shapes occurs: the sequence preceding this moment is reciprocal to that observed after it. A moment when the celestial equator is intersected by a celestial body is a particular point in the histogram series. Its position on the time scale can be determined, with a high accuracy, from the palindrome effect;
4. Changes of histograms near the "equinoxes" depend on the direction in which the celestial body intersects the equator plane (from the northern or from southern hemisphere);
5. Seeming contradiction between these two phenomena — the similarity of the anti-parallel sequences obtained at successive equinoxes, and the mirror similarity of two halves of each histogram sequence (the "palindrome effect") — is, probably, resolved by the fact that these two kinds of symmetry are determined by different, not overlapping sets of similar histograms;
6. Phenomena observed with collimators can be considered as an indication of the anisotropy of space. Vanishing of the "palindrome" phenomena when the collimator is oriented towards the Polar Star is, perhaps, the most indicative fact favoring the conclusion on the anisotropy. Analysis of this phenomenon should be subjected to further studies.

B. V. Komberg attracted our attention to the fact that the direction of the equinox line (the line of intersection of the equatorial and ecliptic planes) may coincide with direction of the line of the minimum temperature of the Cosmic Microwave Backgroud (relict) Radiation, called the "axis of evil" in the scientific literature [7].

Acknowledgements

We sincerely appreciate A. E. Shnoll and A. E. Rodin's participation in the calculation of the moments the celestial equator is intersected by the Moon, Venus and Mars and valuable advice in the data analysis. We sincerely appreciate valuable discussion with S. A. Vasiliev and his interpretation of our results [8]. We sincerely appreciate A. A. Andreyeva's work as a "second expert" in the analysis of similarity of histograms. We are much obliged to our colleagues from the Laboratory for Physical Biochemistry of the Institute for Theoretical and Experimental Biophysics of RAS and from Chair of Biophysics at Physics Department of MSU, for useful discussions at our seminars, and especially to O. Yu. Seraya for discussion and English translation of the article. We appreciate T. A. Zenchenko's systematic job on the development and support of a computer databank for radioactivity monitoring. We are grateful to D. D. Rabounski for his valuable discussion. The work would be impossible without M. N. Kondrashova and her continuous attention, support and discussion of results. Financial support from D. B. Zimin is greatly appreciated.

Submitted on January 01, 2011 / Accepted on January 05, 2011

References

1. Shnoll S. E. Cosmo-Physical Factors in Stochastic Processes. Svenska Fisikarkivet, Stockholm, 2009.
2. Shnoll S. E., Panchelyuga V. A., and Shnoll A. E. The Palindrome Effect. *Progress in Physics*, 2008, v. 2, 151–153.
3. Shnoll S. E. and Rubinstein I. A. Regular changes in the fine structure of histograms revealed in the experiments with collimators which isolate beams of alpha-particles flying at certain directions. *Progress in Physics*, 2009, v. 2, 83–95.
4. Astronomical Yearbook. St. Petersburg, Nauka, 2000–2008.
5. Shnoll S. E. The "scattering of the results of measurements" of processes of diverse nature is determined by the Earth's motion in the inhomogeneous space-time continuum. The effect of "half-year palindromes. *Progress in Physics*, 2009, v. 1, 3–7.
6. Shnoll S. E., Rubinstein I. A., Vedenkin N. N. The "arrow of time" in the experiments in which alpha-activity was measured using collimators directed East and West. *Progress in Physics*, 2010, v. 1, 26–29.
7. Schild R. E., Gibson C. H. Goodness in the Axis of Evil. arXiv: astro-ph/0802.3229.
8. Vasiliev S. A. On the physical model of the phenomena registered in the experiments by Shnoll's group and Smirnov's group. *Progress in Physics*, 2009, v. 2, 29–43.

Electron Configuration, and Element No.155 of the Periodic Table of Elements

Albert Khazan

E-mail: albkhazan@gmail.com

Blocks of the Electron Configuration in the atom are considered with taking into account that the electron configuration should cover also element No.155. It is shown that the electron configuration formula of element No.155, in its graphical representation, completely satisfies Gaussian curve.

1 Introduction

As is known, even the simpliests atoms are very complicate systems. In the centre of such a system, a massive nucleus is located. It consists of protons, the positively charged particles, and neutrons, which are charge-free. Masses of protons and neutrons are almost the same. Such a particle is almost two thousand times heavier than the electron. Charges of the proton and the electron are opposite, but the same in the absolute value. The proton and the neutron differ from the viewpoint on electromagnetic interactions. However in the scale of atomic nuclei they does not differ. The electron, the proton, and the neutron are subatomic articles. The theoretical physicists still cannot solve Schrödinger's equation for the atoms containing two and more electrons. Therefore, they process the calculations for only the single-electron atom of hydrogen, with use of the dualistic property of the electron, according to which it can be respresented, equally, as a particle and a wave. At the same time, the conclusions provided after the quantum theory cannot be considered as the finally true result.

To make the further text simpler, we assume the following brief notations: the Periodic Table of Elements containing 118, 168, and 218 elements will be referred to as T.118, T.168, and T.218 respectively.

2 Calculation of the electron shell for element No.155

Electron shells of the atoms (known also as the *levels*) are regularly denoted as K, L, M, N, O, or as plain numbers from 1 to 5. Each level consists of numerous sub-levels, which are split into atomic orbitales. For instance, the 1st level K consists of a single sub-level 1s. The second level L consists of two sub-levels 2s and 2p. The third level M consists of the 3s, 3p, and 3d sub-levels. The fourth level N consists of the 4s, 4p, 4d, and 4f sub-levels. At the same distance from the atomic nucleus, only the following orbitales can exist: one -s-, three -p-, five -d-, seven -f-, while no more than two electrons can be located in each single orbital (according to Pauli's principle). Hence, the number of electrons in each level can be calculated according to the formula $2N^2$. Results of the calculation are given in Table 1.

As is seen from this Table, the complete external electron level is the configuration s2+p6, known as octet.

	K	L	M	N	O	Sum	Content in the shells
s	2					2	in each shell
p	2	6				8	in each, commencing in the 2nd shell
d	2	6	10			18	in each, commencing in the 3rd shell
f	2	6	10	14		32	in each, commencing in the 4th shell
g	2	6	10	14	18	50	in each, commencing in the 5th shell

Table 1: Number of electrons in each level.

The elements, whose electrons occupy the respective sublevels, have one of the denotations: s-, p-, d-, f-, or g-elements (in analogy to electrons).

2.1 Electron Configuration in the other elements

In the regular form of the Periodic Table of Elements, each cell of the Table bears a large information about the element, including the electron constitution of the atom. The cells containing the same sub-levels are often the same-coloured in the Table, and are joined into the following blocks (T.118):

s-elements, the 1st and the 2nd groups, 7 periods;

p-elements, 6 groups × 6 periods (periods 5–10, 13–18, 31–36, 49–54, 81–86, 113–118);

d-elements, 10 groups × 4 periods, between s- and p-elements;

f-elements, 2 lines of 14 elements each (lantanides and actinides).

Fig. 1 shows distribution of the blocks of T.118, with the assumption of that all last elements are known (the lower arc) [1]. The tabular data of the blocks are easy-to-convert into a graph, if using the known number of the elements. It should be noted that the abscissa axis means number of the blocks (not number of the periods). The form of this arc is close to parabola, and is easy-to-describe by the cubic equation with the value of true approximation $R^2 = 1$.

One can find, in the scientific press, suggestions about the possibility of introducing, into the version T.118 of the Periodic Table, two additional periods of 50 elements in each thus

making it T.218. Therefore, we checked this variant as well (the upper arc), for clarity of the experiments [2, 3]. According to the reference data [4], we assumed five blocks which join all elements of the Periodic Table as follows:

s-elements = 18,
p-elements = 48,
d-elements = 60,
f-elements = 56,
g-elements = 36.

As is seen, the upper arc in Fig. 1 is absolutely similar to that of T.118 (the lower arc). The larger size of the upper arc (T.218) are due to the larger number of elements.

Having these two examples considered, we clearly understand that the aforeapplied method we suggested can as well be applied to the version of the Periodic Table which ends at element No.155.

In order to check this supposition, we created Table 2 wherein we present the respective data for Fig. 1 and Fig. 2.

The upper arc of Fig. 2 shows distribution of the blocks of the electron configuration, calculated according to the reference data of T.168. Lower, another arc is presented. It is created according to our calculation for T.155 (i.e. for the Table of Elements, whose upper limit if element No.155). As is seen, the left branches of the arcs differ from each other for a little, while the right branches actually met each other. The absence of any bends or breaks, and also smooth form of both arcs, and their complete satisfying the approximation equation $R^2 = \sim 1$, manifests the presence of the same law in the basis of these data.

Therefore, we now can claim that element No.155 is included into the blocks of the electron configuration as the last element of the Periodic Table of Elements.

2.2 Electron shells of the atoms

Because our method of comparing the electron configuration of the elements was successful for element No.155, we are going to apply it to theoretical constructing the electron shells. Here we should take into account that: the electrons of the external shells bear more powerful energy, they are more distantly located from the nucleus, and determine the chemical properties of reactions due to the fact that their connexion with the nucleus is weaker thus easier to break. All data, we collected in order to check the aforementioned suggestion, are presented in Table 3. Line 4 of the Table contains the data for the version of element No.155 as that continuing the Table of Elements, while Line 5 contains the respective data suggested by me according to [5].

As is seen, from Fig. 3, all the arcs have the form which is very close to parabola, with a clearly observed maximum and the joined left branches. The difference in their ordinates is due to the difference in the number of the electrons (column 5 of Table 3). The right branches are parallel to each other, and are shifted with respect to each other for the shell number. The main result means here the presence of a qualitative connexion between the electron shells and their graphical representation. For only this reason, we had the possibility to compare the data of the last lines of Table 3.

Fig. 4 manifests that the upper arc is similar to the previous of Fig. 3, while the lower arc (T.155 Author) very differs from all them. According to its form, this is a differential function of normal distribution (the Gauss arc). The difference between the ends of the left and right branches is 0.645%. The branches are very symmetric to each other with respect to the vertical axis coming through the top with coordinates (5, 36). Hence, here is also a strong dependency between the regular method of description of the electron shells and its graphical representation.

This fact is most illustrative manifested in Fig. 5. The left straight covers four electron shells (2, 8, 18, 32), which are the same for all versions of Table 3 (as follows from the equation of the straight line $Y = 2X + 0.6931$. As is seen, once the arcs reach their maximum, they come down very fast (this is because the number of electrons decreases very fast in the shells).

3 Conclusion

Thus, element No.155 has really lawful to be positioned in the Periodic Table of Elements. This element points out not only the upper limit of the Table, found in my earlier study on the basis of the hyperbolic law [6, 7], but also can be presented as a graphical sequel of the calculations produced according to Quantum Mechanics (they have a high precision).

Submitted on February 01, 2011 / Accepted on March 08, 2011

References

1. http://webelements.com; See also: http://en.wikipedia.org/wiki/File:Electron_Configuration_Table.jpg
2. Seaborg G. T. and Bloom J. L. The synthetic elements. *Scientific American*, 1969, v. 220, no. 4, 56.
3. Goldanskii V. I. About connections nuclear and chemical physics. *Progress in Physical Sciences*, 1976, v. 118, issue 2.
4. Nebegrall W. H., Schmidt F. C., Holtzclaw H. F., Jr. General Chemistry. Fourth edition, D. C. Heath and Company, Massachusetts, 1972, pp. 668–670; Also: Extended Periodic Table (suggested by Jeries A. Rihani in 1984), http://jeries.rihani.com
5. FLW, Inc. provides ISO certified calibration for physical measurement, test and control, http://www.flw.com/datatools/periodic/
6. Khazan A. Upper limit in the Periodic System of Elements. *Progress in Physics*, 2007, v. 1, 38–41.
7. Khazan A. Upper Limit in Mendeleev's Periodic Table — Element No.155. Svenska fysikarkivet, Stockholm, 2010.

Number of the elements	Number of the blocks				
	s	p	d	f	g
T.218	18	48	60	56	36
T.168	16	42	50	42	18
T.118	14	36	40	28	—
T.155	16	36	46	42	15

Table 2: Blocks of the electron configuration.

Number of the elements	Number of the electrons in the shells								
T.218	2	8	18	32	50	50	32	18	8
T.168	2	8	18	32	50	32	18	8	—
T.118	2	8	18	32	32	18	8	—	—
T.155 Table	2	8	18	32	50	32	11	2	—
T.155 Author	2	8	18	32	36	32	18	8	1

Table 3: Electron shells of the atoms.

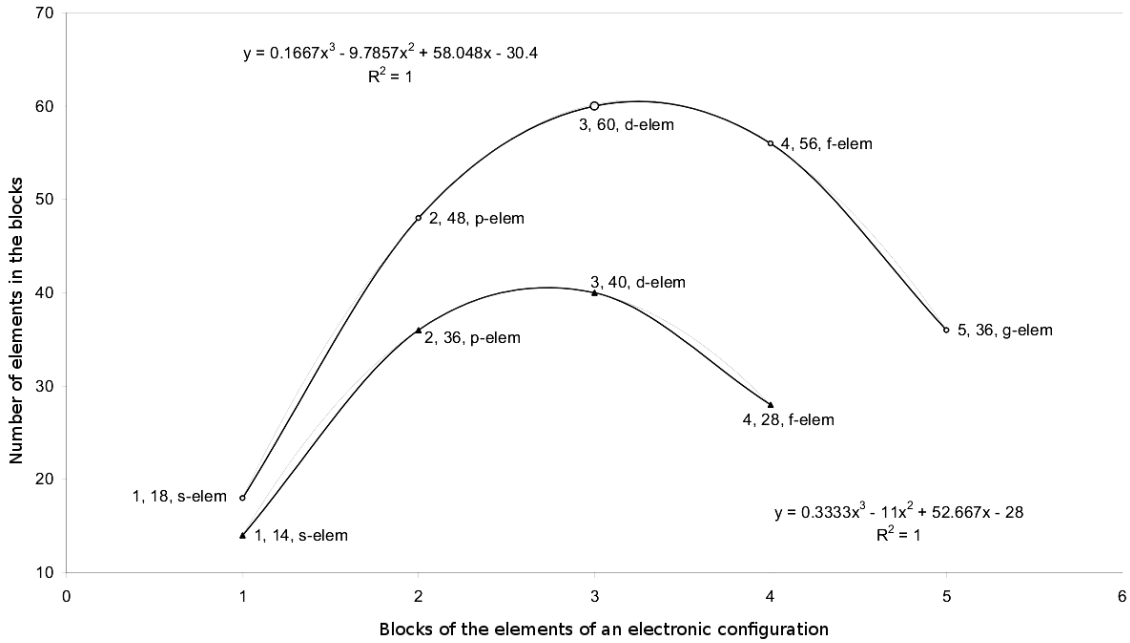


Fig. 1: Location of the blocks of the electron configuration in the Periodic Table of Elements, containing different number of the elements. The upper arc — the Table of 218 elements. The lower arc — the Table of 118 elements.

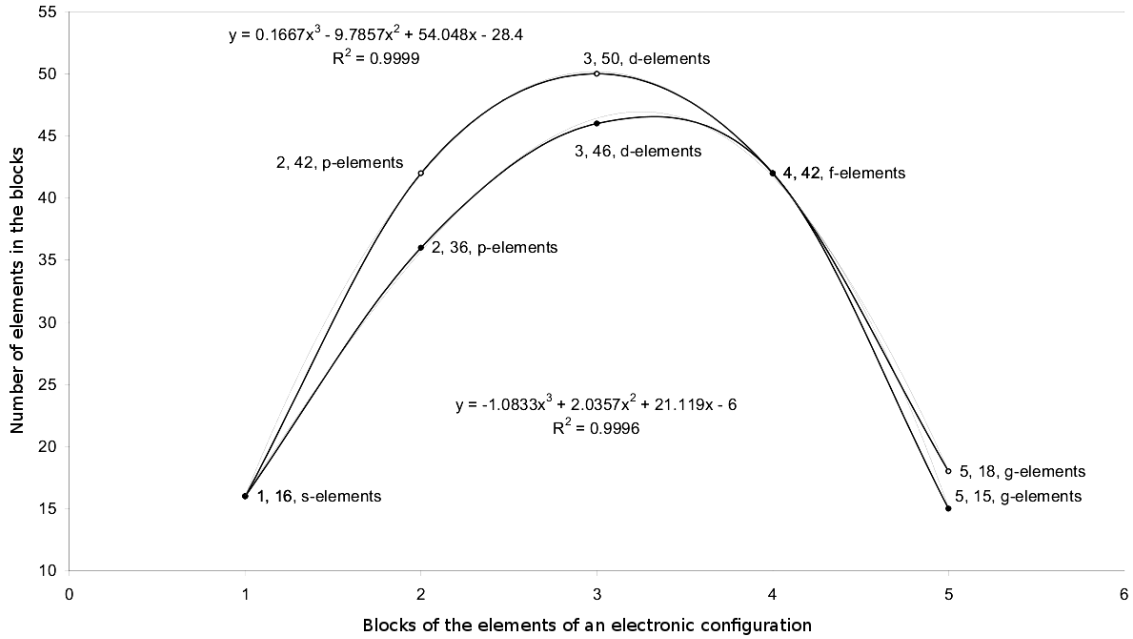


Fig. 2: Dependency, in the blocks, between the number of the elements and the electron configuration. The upper arc — the Table of 168 elements. The lower arc — the Table of 155 elements.

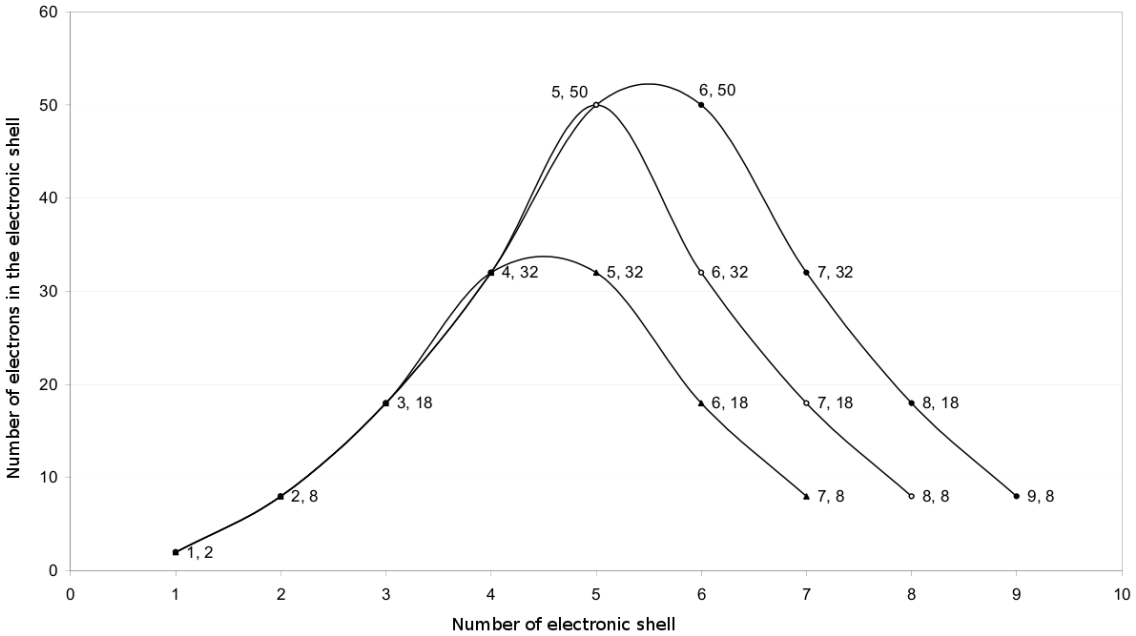


Fig. 3: Dependency of the number of electrons in the electron shells from the shell number, for three versions of the Periodic Table of Elements — T.118, T.168, T.218 (from up to down).

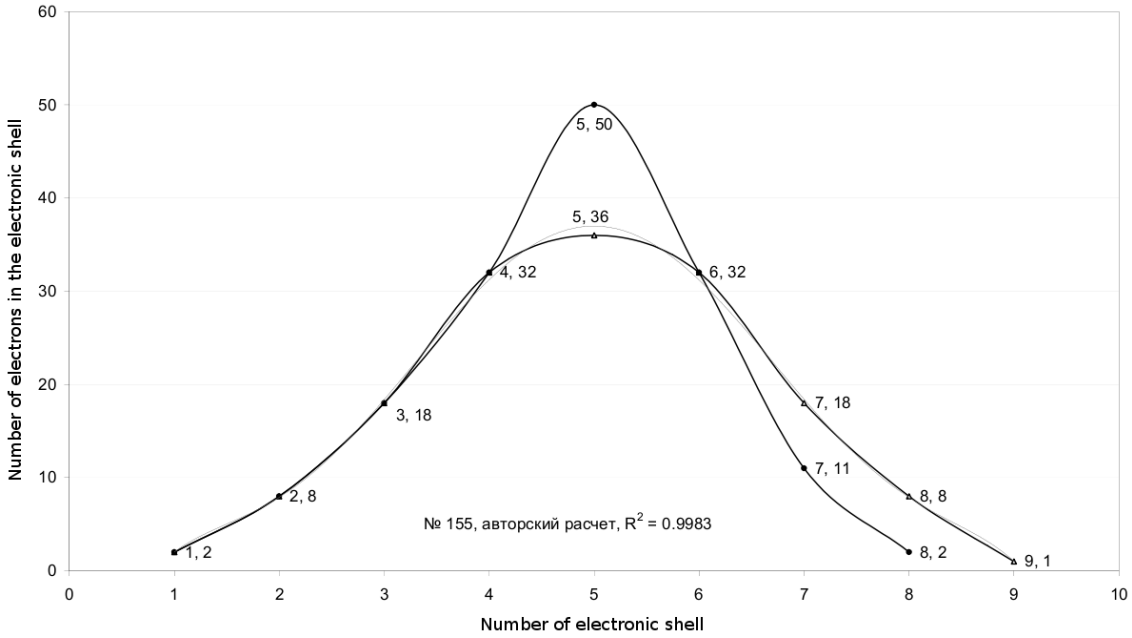


Fig. 4: Dependency of the number of electrons in the electron shells from the shell number, for element No.155 according to the tabular data (the upper arc) and the author's calculation (the lower arc).

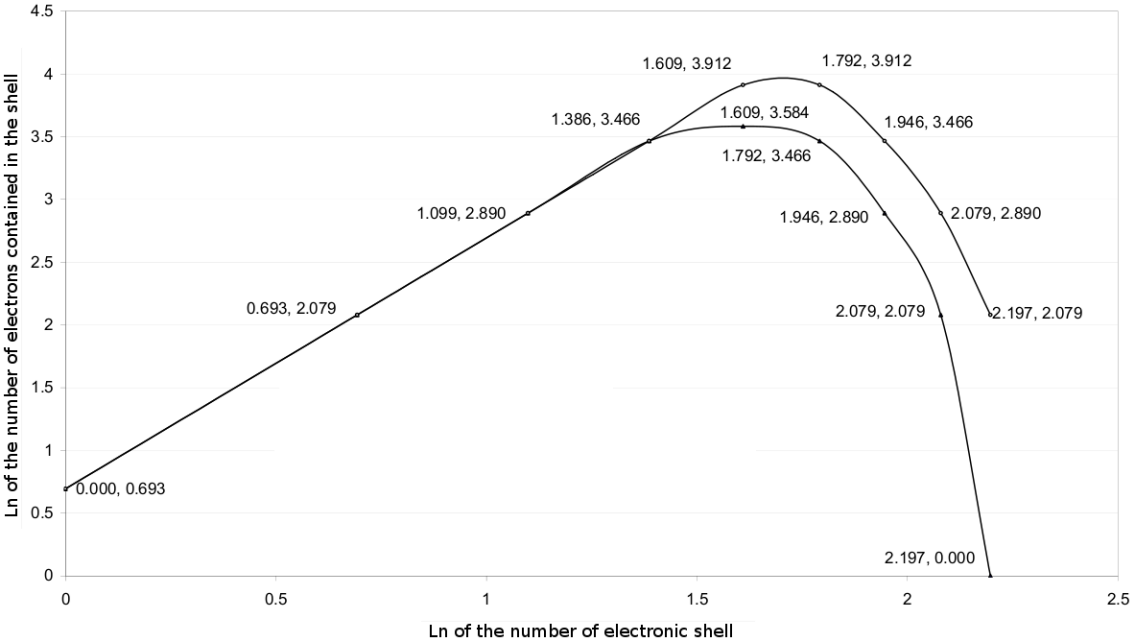


Fig. 5: Dependency of the number of electrons in the electron shells from the shell number (presented in the logarithm coordinates), for T.218 (the upper arc) and for T.155 according to the author's calculation (the lower arc).

Dynamical 3-Space: Cosmic Filaments, Sheets and Voids

Reginald T. Cahill
School of Chemical and Physical Sciences, Flinders University, Adelaide 5001, Australia
E-mail: Reg.Cahill@flinders.edu.au

Observations of weak gravitational lensing combined with statistical tomographic techniques have revealed that galaxies have formed along filaments, essentially one-dimensional lines or strings, which form sheets and voids. These have, in the main, been interpreted as "dark matter" effects. To the contrary here we report the discovery that the dynamical 3-space theory possesses such filamentary solutions. These solutions are purely space self-interaction effects, and are attractive to matter, and as well generate electromagnetic lensing. This theory of space has explained bore hole anomalies, supermassive black hole masses in spherical galaxies and globular clusters, flat rotation curves of spiral galaxies, and other gravitational anomalies. The theory has two constants, G and α, where the bore hole experiments show that $\alpha \approx 1/137$ is the fine structure constant.

1 Introduction

Observations of weak gravitational lensing and statistical tomographic techniques have revealed that galaxies have formed along filaments, essentially one-dimensional lines or strings [1], see Fig.1. These have, in the main, been interpreted as "dark matter" effects. Here we report the discovery that the dynamical 3-space theory possesses such filamentary solutions, and so does away with the "dark matter" interpretation. The dynamical 3-space theory is a uniquely determined generalisation of Newtonian gravity, when that is expressed in terms of a velocity field, instead of the original gravitational acceleration field [2, 3]. This velocity field has been repeatedly detected via numerous light speed anisotropy experiments, beginning with the 1887 Michelson-Morley gas-mode interferometer experiment [4, 5]. This is a theory of space, and has explained bore hole anomalies, supermassive black hole masses in spherical galaxies and globular clusters, flat rotation curves of spiral galaxies, and other gravitational anomalies. The theory has two constants, G and α, where the bore hole experiments show that $\alpha \approx 1/137$ is the fine structure constant. The filamentary solutions are purely a consequence of the space self-interaction dynamics, and are attractive to matter, and as well generate electromagnetic lensing. The same self-interaction dynamics has been shown to generate inflow singularities, *viz* black holes [6], with both the filaments and black holes generating long-range non-Newtonian gravitational forces. The dynamical 3-space also has Hubble expanding universe solutions that give a parameter-free account of the supernova redshift-magnitude data, without the need for "dark matter" or "dark energy" [7]. The black hole and filament solutions are primordial remnants of the big bang in the epoch when space was self-organising, and then provided a framework for the precocious clumping of matter, as these inflow singularities are long-range gravitational attractors. That α determines the strength of these phenomena implies that we are seeing evidence of a unification of space, gravity and quantum theory, as conjectured in Process Physics [2].

2 Dynamical 3-Space

The dynamics of space is easily determined by returning to Galileo's discoveries of the free-fall acceleration of test masses, and using a velocity field to construct a minimal and unique formulation that determines the acceleration of space itself [2, 8]. In the case of zero vorticity we find

$$\nabla\cdot\left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}\right) + \frac{\alpha}{8}\left((trD)^2 - tr(D^2)\right) + ... = -4\pi G\rho \quad (1)$$

$$\nabla\times\mathbf{v} = \mathbf{0}, \quad D_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right), \quad (2)$$

G is Newton's constant, which has been revealed as determining the dissipative flow of space into matter, and α is a dimensionless constant, that experiment reveals to be the fine structure constant. The space acceleration is determined by the Euler constituent acceleration

$$\mathbf{a} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} \quad (3)$$

The matter acceleration is found by determining the trajectory of a quantum matter wavepacket to be [9]

$$\mathbf{g} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} + (\nabla\times\mathbf{v})\times\mathbf{v}_R \quad (4)$$

$$-\frac{\mathbf{v}_R}{1-\dfrac{\mathbf{v}_R^2}{c^2}}\frac{1}{2}\frac{d}{dt}\left(\frac{\mathbf{v}_R^2}{c^2}\right) + ... \quad (5)$$

where $\mathbf{v}(\mathbf{r}, t)$ is the velocity of a structured element of space wrt to an observer's arbitrary Euclidean coordinate system, but which has no ontological meaning. The relativistic term in (5) follows from extremising the elapsed proper time wrt

a quantum matter wave-packet trajectory $\mathbf{r}_o(t)$, see [2]. This ensures that quantum waves propagating along neighbouring paths are in phase.

$$\tau = \int dt \sqrt{1 - \frac{\mathbf{v}_R^2(\mathbf{r}_0(t), t)}{c^2}} \qquad (6)$$

where $\mathbf{v}_R(\mathbf{r}_o(t), t) = \mathbf{v}_o(t) - \mathbf{v}(\mathbf{r}_o(t), t)$, is the velocity of the wave packet, at position $\mathbf{r}_0(t)$, wrt the local 3-space, and $\mathbf{g} = d\mathbf{r}_O/dt$. This shows that (i) the matter "gravitational" geodesic is a quantum wave refraction effect, with the trajectory determined by a Fermat maximum proper-time principle, and (ii) that quantum systems undergo a local time dilation effect caused by their absolute motion wrt space. The last term in (5) causes the precession of planetary orbits.

It is essential that we briefly review some of the many tests that have been applied to this dynamical 3-space.

2.1 Direct Observation of 3-Space

Numerous direct observations of 3-space involve the detection of light speed anisotropy. These began with the 1887 Michelson-Morley gas-mode interferometer experiment, that gives a solar system galactic speed in excess of 300 km/s, [4, 5]*. These experiments have revealed components of the flow, a dissipative inflow, caused by the sun and the earth, as well as the orbital motion of the earth. The largest effect is the galactic velocity of the solar system of 486 km/s in the direction RA = 4.3°, Dec = −75°, determined from spacecraft earth-flyby Doppler shift data [10], a direction first detected by Miller in his 1925/26 gas-mode Michelson interferometer experiment [11].

2.2 Newtonian Gravity Limit

In the limit of zero vorticity and neglecting relativistic effects (2) and (5) give

$$\nabla \cdot \mathbf{g} = -4\pi G\rho - 4\pi G\rho_{DM}, \quad \nabla \times \mathbf{g} = \mathbf{0} \qquad (7)$$

where

$$\rho_{DM} = \frac{\alpha}{32\pi G}\left((trD)^2 - tr(D^2)\right). \qquad (8)$$

This is Newtonian gravity, but with the extra dynamical term which has been used to define an effective "dark matter" density. This is not necessarily non-negative, so in some circumstances ant-gravity effects are possible, though not discussed herein.This ρ_{DM} is not a real matter density, of any form, but is the matter density needed within Newtonian gravity to explain dynamical effects caused by the α-term in (2). This term explains the flat rotation curves of spiral galaxies, large light bending and lensing effects from galaxies, and other effects. However, it is purely a space self-interaction effect.

*Amazingly it continues to be claimed that this experiment was null.

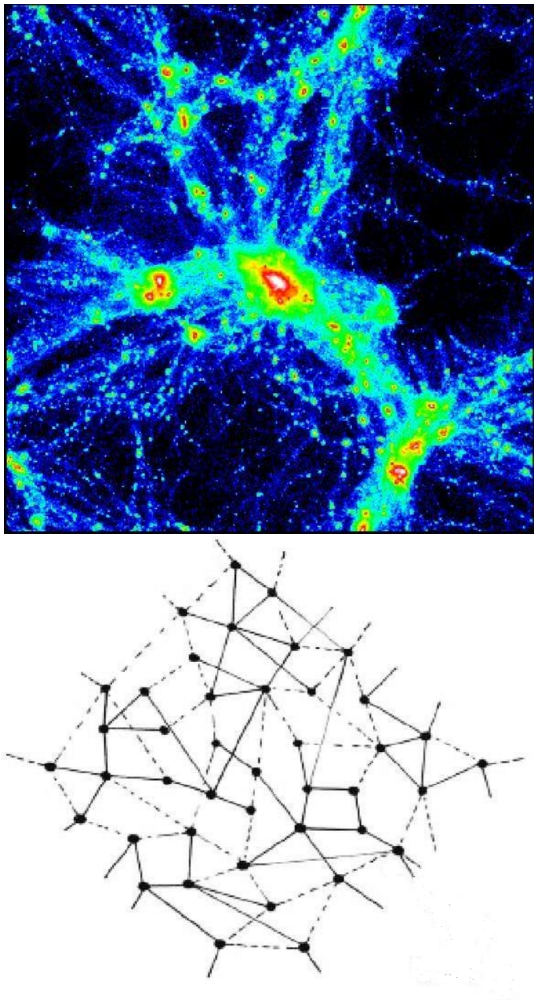


Fig. 1: Top: Cosmic filaments as revealed by gravitational lensing statistical tomography. From J.A. Tyson and G. Bernstein, Bell Laboratories, Physical Sciences Research, http://www.bell-labs.com/org/physicalsciences/projects/darkmatter/darkmatter.html. Bottom: Cosmic network of primordial filaments and primordial black holes, as solution from (2).

2.3 Curved Spacetime Formalism

Eqn.(6) for the elapsed proper time may be written

$$d\tau^2 = dt^2 - \frac{1}{c^2}(d\mathbf{r}(t) - \mathbf{v}(\mathbf{r}(t), t)dt)^2 = g_{\mu\nu}(x)dx^\mu dx^\nu, \qquad (9)$$

which introduces a curved spacetime metric $g_{\mu\nu}$. However this spacetime has no ontological significance — it is merely a mathematical artifact, and as such hides the underlying dynamical 3-space. Its only role is to describe the geodesic of the matter quantum wave-packet in gerneral coordinates. The metric is determined by solutions of (2). This induced metric is not determined by the Einstein-Hilbert equations, which originated as a generalisation of Newtonian gravity, but without the knowledge that a dynamical 3-space had indeed been

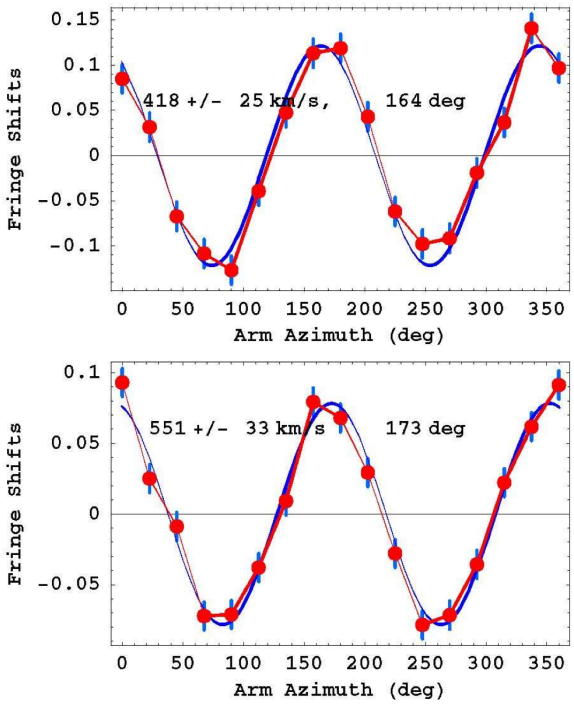


Fig. 2: (a) A typical Miller averaged-data from September 16, 1925, $4^h40'$ Local Sidereal Time (LST) — an average of data from 20 turns of the gas-mode Michelson interferometer. Plot and data after fitting and then subtracting both the temperature drift and Hicks effects from both, leaving the expected sinusoidal form. The error bars are determined as the rms error in this fitting procedure, and show how exceptionally small were the errors, and which agree with Miller's claim for the errors. (b) Best result from the Michelson-Morley 1887 data — an average of 6 turns, at 7^h LST on July 11, 1887. Again the rms error is remarkably small. In both cases the indicated speed is v_P — the 3-space speed projected onto the plane of the interferometer. The angle is the azimuth of the 3-space speed projection at the particular LST. The speed fluctuations from day to day significantly exceed these errors, and reveal the existence of 3-space flow turbulence — i.e. gravitational waves.

detected by Michelson and Morley in 1887 by detecting light speed anisotropy.

2.4 Gravitational Waves

Eqn.(2) predicts time dependent flows, and these have been repeatedly detected, beginning with the Michelson and Morley experiment in 1887. Apart from the sidereal earth-rotation induced time-dependence, the light-speed anisotropy data has always shown time-dependent fluctuations/turbulence, and at a scale of some 10% of the background galactic flow speed. This time dependent velocity field induces "ripples" in the spacetime metric in (9), which are known as "gravitational waves". They cannot be detected by a vacuum-mode Michelson interferometer.

2.5 Matter Induced Minimal Black Holes

For the special case of a spherically symmetric flow we set $\mathbf{v}(\mathbf{r}, t) = -\hat{\mathbf{r}}v(r, t)$. Then (2) becomes, with $v' = \partial v/\partial r$,

$$\frac{\partial v'}{\partial t} + vv'' + \frac{2}{r}vv' + (v')^2 + \frac{\alpha}{2r}\left(\frac{v^2}{2r} + vv'\right) = -4\pi G\rho \qquad (10)$$

For a matter density $\rho(r)$, with maximum radius R, (10) has an exact inhomogeneous static solution [12]

$$v(r)^2 = \begin{cases} \dfrac{2G}{(1-\frac{\alpha}{2})r}\displaystyle\int_0^r 4\pi s^2\rho(s)ds \\ +\dfrac{2G}{(1-\frac{\alpha}{2})r^{\frac{\alpha}{2}}}\displaystyle\int_r^R 4\pi s^{1+\frac{\alpha}{2}}\rho(s)ds, & 0<r\leq R \\ \dfrac{2\gamma}{r}, \qquad r>R \end{cases} \qquad (11)$$

where

$$\gamma = \frac{G}{(1-\frac{\alpha}{2})}\int_0^R 4\pi s^2\rho(s)ds = \frac{GM}{(1-\frac{\alpha}{2})} \qquad (12)$$

Here M is the total matter mass. As well the middle term in (11) also has a $1/r^{\alpha/2}$ inflow-singularity, but whose strength is mandated by the matter density, and is absent when $\rho(r) = 0$ everywhere. This is a minimal attractor or "black hole"*, and is present in all matter systems. For the region outside the sun, $r > R$, Keplerian orbits are known to well describe the motion of the planets within the solar system, apart from some small corrections, such as the Precession of the Perihelion of Mercury, which follow from relativistic term in (2). The sun, as well as the earth, has only an induced "minimal attractor", which affects the interior density, temperature and pressure profiles [12]. These minimal black holes contribute to the external $g = GM^\star/r^2$ gravitational acceleration, through an effective mass $M^\star = M/(1-\alpha/2)$. The 3-space dynamics contributes an effective mass [2]

$$M_{BH} = \frac{M}{1-\frac{\alpha}{2}} - M = \frac{\alpha}{2}\frac{M}{1-\frac{\alpha}{2}} \approx \frac{\alpha}{2}M. \qquad (13)$$

These induced black hole "effective" masses have been detected in numerous globular clusters and spherical galaxies and their predicted effective masses have been confirmed in some 19 such cases, as shown in Fig. 3, [6]. The non-Newtonian effects in (11) are also detectable in bore hole experiments.

2.6 Earth Bore Holes Determine α

The value of the parameter α in (2) was first determined from earth bore hole g-anomaly data, which shows that gravity decreases more slowly down a bore hole than predicted by Newtonian gravity, see Figs.4 and 5. From (5) and (11) we find

*The term "black hole" refers to the existence of an event horizon, where the in-flow speed reaches c, but otherwise has no connection to the putative "black holes" of GR.

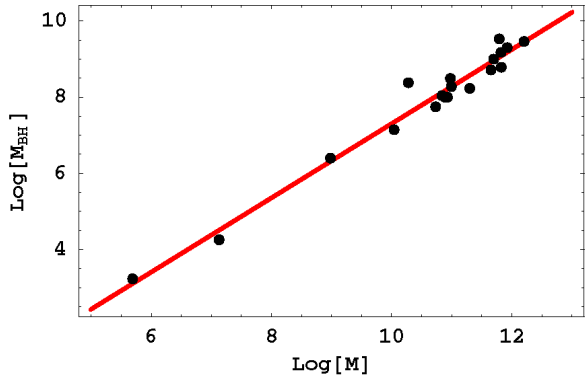


Fig. 3: The data shows $\mathrm{Log}_{10}[M_{BH}]$ for the minimal induced black hole masses M_{BH} for a variety of spherical matter systems, from Milky Way globular clusters to spherical galaxies, with masses M, plotted against $\mathrm{Log}_{10}[M]$, in solar masses M_0. The straight line is the prediction from (13) with $\alpha = 1/137$. See [6] for references to the data.

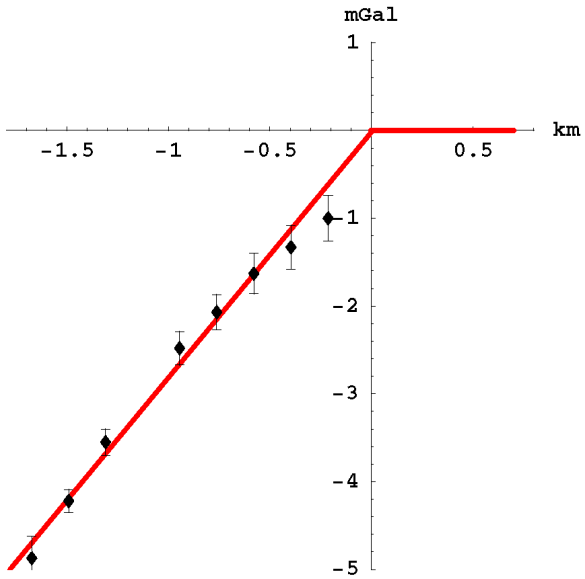


Fig. 4: The data shows the gravity residuals for the Greenland Ice Shelf [13] Airy measurements of the $g(r)$ profile, defined as $\Delta g(r) = g_{Newton} - g_{observed}$, and measured in mGal (1mGal = 10^{-3} cm/s^2) and plotted against depth in km. The borehole effect is that Newtonian gravity and the new theory differ only beneath the surface, provided that the measured above-surface gravity gradient is used in both theories. This then gives the horizontal line above the surface. Using (15) we obtain $\alpha^{-1} = 137.9 \pm 5$ from fitting the slope of the data, as shown. The non-linearity in the data arises from modelling corrections for the gravity effects of the irregular sub ice-shelf rock topography. The ice density is 920 kg/m^3. The near surface data shows that the density of the Greenland ice, compressed snow, does not reach its full density until some 250m beneath the surface — a known effect.

the gravitational acceleration at radius $r = R + d$ to be

$$g(d)=\begin{cases} -\dfrac{GM}{(1-\alpha/2)(R+d)^2}+\dfrac{2\pi G\rho(R)d}{(1-\alpha/2)}+\ldots \\ \qquad -\dfrac{4\pi R^2 G\rho(R)G}{(1-\alpha/2)(R+d)^2}, \quad d<0 \\ -\dfrac{GM}{(1-\alpha/2)(R+d)^2}, \quad d>0 \end{cases} \tag{14}$$

In practice the acceleration above the earth's surface must be measured in order to calibrate the anomaly, which defines the coefficient $\overline{GM} = GM/(1-\alpha/2)$ in (14). Then the anomaly is

$$\Delta g = g_{NG}(d) - g(d) = 2\pi\alpha G\rho(R)d + O(\alpha^2), \quad d<0 \tag{15}$$

to leading order in α, and where $g_{NG}(d)$ is the Newtonian gravity acceleration, given the value of $\overline{GM}$ from the above-surface calibration, for a near-surface density $\rho(R)$. The experimental data then reveals α to be the fine structure constant, to within experimental errors [6]. The experiments have densities that differ by more than a factor of 2, so the result is robust.

2.7 G Measurement Anomalies

There has been a long history of anomalies in the measurement of Newton's gravitational constant G, see Fig. 7. The explanation is that the gravitational acceleration external to a piece of matter is only given by application of Newton's inverse square law for the case of a spherically symmetric mass. For other shapes the α-dependent interaction in (2) results in forces that differ from Newtonian gravity at $O(\alpha)$. The anomalies shown in Fig. 7 result from analysing the one-parameter, G, Newtonian theory, when gravity requires a two parameter, G and α, analysis of the data. The scatter in the measured G values appear to be of $O(\alpha/4)$. This implies that laboratory measurements to determine G will also measure α [2].

2.8 Expanding Universe

The dynamical 3-space theory (2) has a time dependent expanding universe solution, in the absence of matter, of the Hubble form $v(r,t) = H(t)r$ with $H(t) = 1/(1+\alpha/2)t$, giving a scale factor $a(t) = (t/t_0)^{4/(4+\alpha)}$, predicting essentially a uniform expansion rate. This results in a parameter-free fit to the supernova redshift-magnitude data, as shown in fig.8, once the age $t_0 = 1/H_0$ of the universe at the time of observation is determined from nearby supernova. In sharp contrast the Friedmann model for the universe has a static solution — no expansion, unless there is matter/energy present. However to best fit the supernova data fictitious "dark matter" and "dark energy" must be introduced, resulting in the ΛCDM model. The amounts $\Omega_\Lambda = 0.73$ and $\Omega_{DM} = 0.23$ are eas-

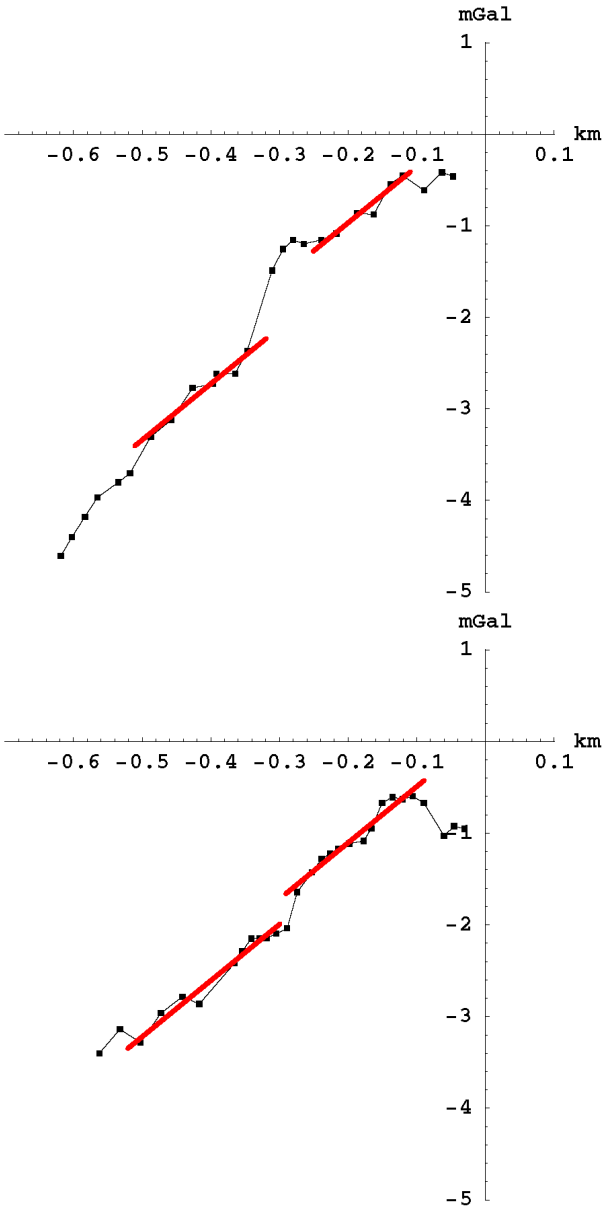


Fig. 5: Gravity residuals $\Delta g(r)$ from two of the Nevada bore hole experiments [14] that give a best fit of $\alpha^{-1} = 136.8 \pm 3$ on using (15). Some layering of the rock is evident. The rock density is 2000 kg/m^3 in the linear regions.

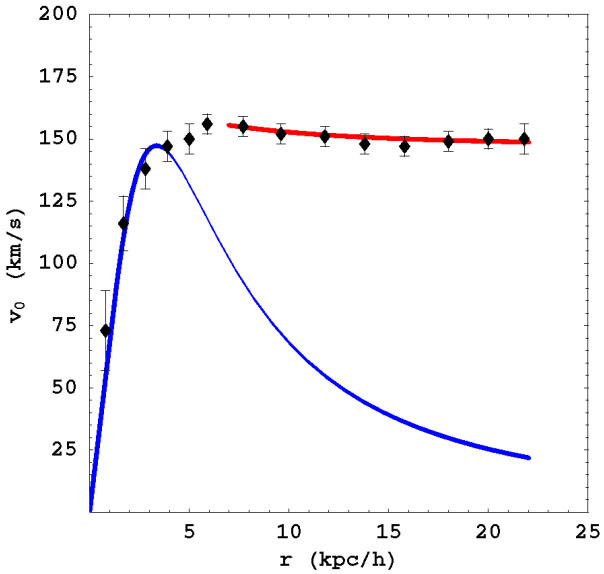


Fig. 6: Plots of the rotation speed data for the spiral galaxy NGC3198. Lower curve shows Newtonian gravity prediction, while upper curve shows asymptotic flat rotation speeds from (19).

ily determined by best fitting the ΛCDM model to the above uniformly expanding result, without reference to the observational supernova data. But then the ΛCDM has a spurious exponential expansion which becomes more pronounced in the future. This is merely a consequence of extending a poor curve fitting procedure beyond the data. The 3-space dynamics (2) results in a hotter universe in the radiation dominated epoch, with effects on Big Bang Nucleosynthesis [15], and also a later decoupling time of some 1.4×10^6 years.

3 Primordial Black Holes

In the absence of matter the dynamical 3-space equation (2) has black hole solutions of the form

$$v(r) = -\frac{\beta}{r^{\alpha/4}} \tag{16}$$

for arbitrary β, but only when $\alpha \neq 0$. This will produce a long range gravitational acceleration, essentially decreasing like $1/r$,

$$g(r) = -\frac{\alpha\beta^2}{4r^{1+\alpha/2}} \tag{17}$$

as observed in spiral galaxies. The inflow in (16) describes an inflow singularity or "black hole" with arbitrary strength. This is unrelated to the putative black holes of General Relativity. This corresponds to a primordial black hole. The dark matter density for these black holes is

$$\rho_{DM}(r) = \frac{\alpha\beta^2(2-\alpha)}{256\pi G r^{2+\alpha/2}} \tag{18}$$

This decreases like $1/r^2$ as indeed determined by the "dark matter" interpretation of the flat rotation curves of spiral galaxies. Here, however, it is a purely 3-space self-interaction effect.

In general a spherically symmetric matter distribution may have a static solution which is a linear combination of the inhomogeneous matter induced solution in (11) and the square of the homogeneous primordial black hole solution in (16), as (10) is linear in $v(r)^2$ and its spatial derivatives.

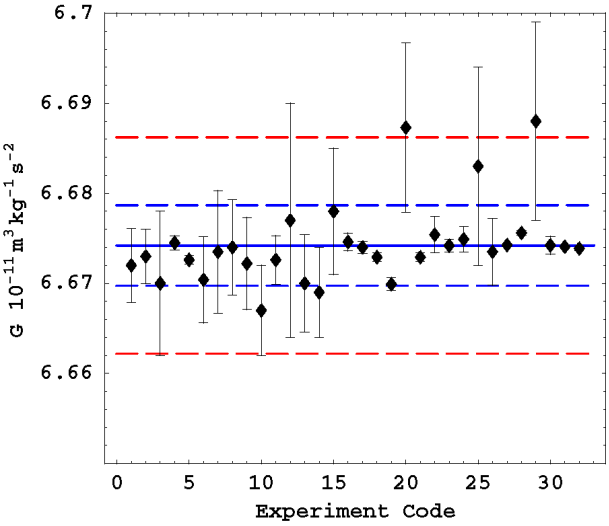


Fig. 7: Results of precision measurements of G published in the last sixty years in which the Newtonian theory was used to analyse the data. These results show the presence of a systematic effect, not in the Newtonian theory, of fractional size $\Delta G/G \approx \alpha/4$. The upper horizontal dashed line shows the value of G from ocean Airy measurements [17], while the solid line shows the current CODATA G value of $6.67428(\pm 0.00067) \times 10^{-11} m^3/kgs^2$, with much lager experimental data range, exceeding $\pm\alpha G/8$, shown by dashed lines as a guide. The lower horizontal line shows the actual value of G after removing the space self-interaction effects via $G \to (1 - \alpha/2)G$ from the ocean value of G. The CODATA G value, and its claimed uncertainty, is seen to be spurious.

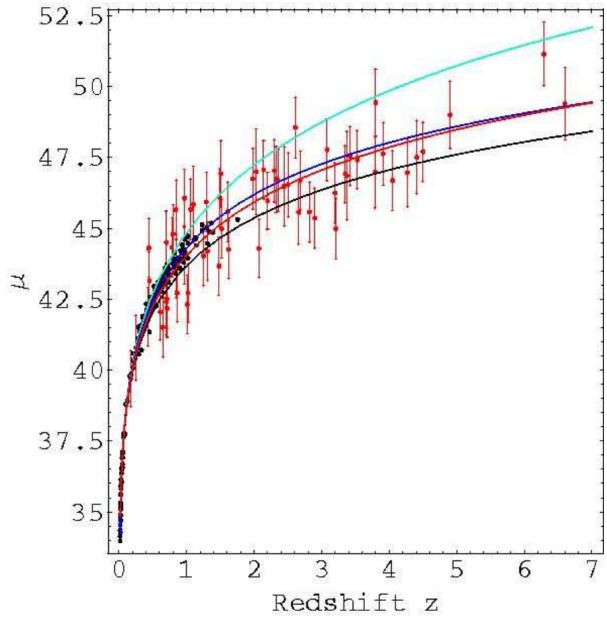


Fig. 8: Hubble diagram showing the supernovae data using several data sets, and the Gamma-Ray-Bursts data (with error bars). Upper curve (green) is ΛCDM "dark energy" only $\Omega_\Lambda = 1$, lower curve (black) is ΛCDM matter only $\Omega_M = 1$. Two middle curves show best-fit of ΛCDM "dark energy"-"dark-matter" (blue) and dynamical 3-space prediction (red), and are essentially indistinguishable. We see that the best-fit ΛCDM "dark energy"-"dark-matter" curve essentially converges on the uniformly-expanding parameter-free dynamical 3-space prediction. The supernova data shows that the universe is undergoing a uniform expansion, wherein a fit to the FRW-GR expansion was forced, requiring "dark energy", "dark matter" and a future "exponentially accelerating expansion".

However this is unlikely to be realised, as a primordial black hole would cause a precocious in-fall of matter, which is unlikely to remain spherically symmetric, forming instead spiral galaxies.

3.1 Spiral Galaxy Rotation Curves

Spiral galaxies are formed by matter in-falling on primordial black hole, leading to rotation of that matter, as the in-fall will never be perfectly symmetric. The black hole acceleration in (17) would support a circular matter orbit with orbital speed

$$v_o(r) = \frac{(\alpha\beta^2)^{1/2}}{2r^{\alpha/4}} \tag{19}$$

which is the observed asymptotic "flat" orbital speed in spiral galaxies, as illustrated in Fig. 6 for the spiral galaxy NGC3198. So the flat rotation curves are simply explained by (2).

4 Primordial Filaments

Eqn.(2) also has cosmic filament solutions. Writing (2) in cylindrical coordinates (r, z, ϕ), and assuming cylindrical symmetry with translation invariance along the z axis, we have for a radial flow $v(r, t)$

$$\frac{1}{r}\frac{\partial v}{\partial t} + \frac{\partial v'}{\partial t} + \frac{vv'}{r} + v'^2 + vv'' + \alpha\frac{vv'}{4r} = 0 \tag{20}$$

where here the radial distance r is the distance perpendicular to the z axis. This has static solutions with the form

$$v(r) = -\frac{\mu}{r^{\alpha/8}} \tag{21}$$

for arbitrary μ. The gravitational acceleration is long-range and attractive to matter, i.e. **g** is directed inwards towards the filament,

$$g(r) = -\frac{\alpha\mu^2}{8r^{1+\alpha/4}} \tag{22}$$

This is for a single infinite-length filament. The dark matter density (8) is

$$\rho_{DM}(r) = -\frac{\alpha\mu^2}{1024\pi G r^{2+\alpha/4}} \tag{23}$$

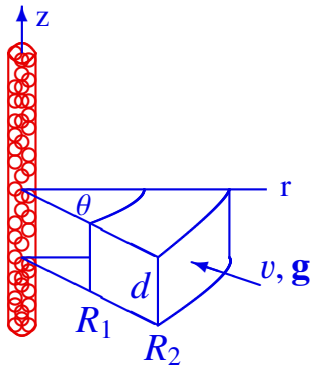


Fig. 9: Sector integration volume, with radii R_1 and R_2, about a filament. For the filament to exist the quantum foam substructure to 3-space must be invoked at short distances.

and negative. But then (7), with $\rho = 0$, would imply a repulsive matter acceleration by the filament, and not attractive as in (22). To resolve this we consider the sector integration volume in Fig.9. We obtain from (22) and using the divergence theorem (in which **dA** is directed outwards from the integration volume)

$$\int_{\mathcal{V}} \nabla\cdot\mathbf{g}dv = \int_{\mathcal{A}} \mathbf{g}\cdot\mathbf{dA} = \frac{\alpha\mu^2\theta d}{8}\left(\frac{1}{R_1^{\alpha/4}} - \frac{1}{R_2^{\alpha/4}}\right) \qquad (24)$$

which is positive because $R_1 < R_2$. This is consistent with (7) for the negative ρ_{DM}, but only if R_1 is finite. However if $R_1 = 0$, as for the case of the integration sector including the filament axis, there is no R_1 term in (24), and the integral is now negative. This implies that (21) cannot be the solution for some small r. The filament solution is then only possible if the dynamical 3-space equation (1) is applicable only to macroscopic distances, and at short distances higher order derivative terms become relevant, such as $\nabla^2(\nabla\cdot\mathbf{v})$. Such terms indicate the dynamics of the underlying quantum foam, with (1) being a derivative expansion, with higher order dervatives becoming more significant at shorter distances.

5 Filament Gravitational Lensing

We must generalise the Maxwell equations so that the electric and magnetic fields are excitations within the dynamical 3-space, and not of the embedding space. The minimal form in the absence of charges and currents is

$$\begin{aligned} \nabla\times\mathbf{E} &= -\mu_0\left(\frac{\partial\mathbf{H}}{\partial t} + \mathbf{v}.\nabla\mathbf{H}\right), \quad \nabla.\mathbf{E} = \mathbf{0}, \\ \nabla\times\mathbf{H} &= \epsilon_0\left(\frac{\partial\mathbf{E}}{\partial t} + \mathbf{v}.\nabla\mathbf{E}\right), \quad \nabla.\mathbf{H} = \mathbf{0} \end{aligned} \qquad (25)$$

which was first suggested by Hertz in 1890 [16], but with $\mathbf{v}$ then being only a constant vector field. As easily determined the speed of EM radiation is now $c = 1/\sqrt{\mu_0\epsilon_0}$ with respect to the 3-space. The time-dependent and inhomogeneous velocity field causes the refraction of EM radiation. This can be computed by using the Fermat least-time approximation. This ensures that EM waves along neighbouring paths are in phase. Then the EM ray paths $\mathbf{r}(t)$ are determined by minimising the elapsed travel time:

$$T = \int_{s_i}^{s_f} \frac{ds|\frac{d\mathbf{r}}{ds}|}{|c\hat{\mathbf{v}}_R(s) + \mathbf{v}(\mathbf{r}(\mathbf{s}),\mathbf{t}(\mathbf{s})|}, \qquad (26)$$

$$\mathbf{v}_R = \frac{d\mathbf{r}}{dt} - \mathbf{v}(\mathbf{r}(\mathbf{t}),\mathbf{t}) \qquad (27)$$

by varying both $\mathbf{r}(s)$ and $t(s)$, finally giving $\mathbf{r}(t)$. Here s is a path parameter, and $c\hat{\mathbf{v}}_R$ is the velocity of the EM radiation wrt the local 3-space, namely c. The denominator in (26) is the speed of the EM radiation wrt the observer's Euclidean spatial coordinates. Eqn.(26) may be used to calculate the gravitational lensing by black holes, filaments and by ordinary matter, using the appropriate 3-space velocity field. Because of the long-range nature of the inflow for black holes and filaments, as in (16) and (21), they produce strong lensing, compared to that for ordinary matter*, and also compared with the putative black holes of GR, for which the in-flow speed decreases like $1/\sqrt{r}$, corresponding to the acceleration field decreasing like $1/r^2$. The EM lensing caused by filaments and black holes is the basis of the stochastic tomographic technique for detecting these primordial 3-space structures.

6 Filament and Black Hole Networks

The dynamical 3-space equation produces analytic solutions for the cases of a single primordial black hole, and a single, infinite length, primordial filament. This is because of the high symmetry of theses cases. However analytic solutions corresponding to a network of finite length filaments joining at black holes, as shown in Fig.1, are not known. For this case numerical solutions will be needed. It is conjectured that the network is a signature of primordial imperfections or defects from the epoch when the 3-space was forming, in the earliest moments of the big bang. It is conjectured that the network of filaments and black holes form a cosmic network of sheets and voids. This would amount to a dynamical breakdown of the translation invariance of space. Other topological defects are what we know as quantum matter [2].

7 Conclusions

The recent discovery that a dynamical 3-space exists has resulted in a comprehensive investigation of the new physics, and which has been checked against numerous experimental and observational data. This data ranges from laboratory Cavendish-type G experiments to the expansion of the universe which, the data clearly shows, is occurring at a uniform rate, except for the earliest epochs. Most significantly

*Eqn:(26) produces the known sun light bending [3].

the dynamics of space involves two parameters: G, Newton's gravitational constant, which determines the rate of dissipative flow of space into matter, and α, which determines the space self-interaction dynamics. That this is the same constant that determines the strength of electromagnetic interactions shows that a deep unification of physics is emerging. It is the α term in the space dynamics that determines almost all of the new phenomena. Most importantly the epicycles of spacetime physics, *viz* dark matter and dark energy, are dispensed with.

Submitted on February 16, 2011 / Accepted on February 21, 2011

References

1. Fabian A. C. (ed.), Clusters and Superclustors of Galaxies, *NATO ASI Series C*, Vol. 366, Kluwer, 1992.
2. Cahill R. T. Process Physics: From Information Theory to Quantum Space and Matter, Nova Science Pub., New York, 2005.
3. Cahill R. T. Dynamical 3-Space: A Review, in: Ether Space-time and Cosmology: New Insights into a Key Physical Medium, Duffy M. and Lévy J. (eds.), *Apeiron*, 2009, v. 16, 135–200.
4. Cahill R. T., Kitto K. Michelson-Morley Experiments Revisited, *Apeiron*, 2003, v. 10, no. 2, 104–117.
5. Cahill R. T. The Michelson and Morley 1887 Experiment and the Discovery of Absolute Motion, *Progress in Physics*, 2005, v. 3, 25–29.
6. Cahill R. T. 3-Space Inflow Theory of Gravity: Boreholes, Blackholes and the Fine Structure Constant, *Progress in Physics*, 2006, v. 2, 9–16.
7. Cahill R. T. Unravelling the Dark Matter — Dark Energy Paradigm, *Apeiron*, 2009, v. 16, no. 3, 323–375.
8. Cahill R. T. Dynamical 3-Space: Emergent Gravity, Should the Laws of Gravitation be Reconsidered?, Múnera H. A. (ed.), Montreal: Apeiron 2011.
9. Cahill R. T. Dynamical Fractal 3-Space and the Generalised Schrödinger Equation: Equivalence Principle and Vorticity Effects, *Progress in Physics*, 2006, v. 1, 27–34.
10. Cahill R. T. Combining NASA/JPL One-Way Optical-Fiber Light-Speed Data with Spacecraft Earth-Flyby Doppler-Shift Data to Characterise 3-Space Flow, *Progress in Physics*, 2009, v. 4, 50–64.
11. Miller D. C. The Ether-Drift Experiment and the Determination of the Absolute Motion of the Earth, *Reviews of Modern Physics*, 1933, v. 5, 203–242.
12. May R. D., Cahill R. T. Dynamical 3-Space Gravity Theory: Effects on Polytropic Solar Models, *Progress in Physics*, 2011, v. 1, 49–54.
13. Ander M. E., Zumberge M. A., Lautzenhiser T., Parker R. L., Aiken C. L. V., Gorman M. R., Nieto M. M., Cooper A. P. R., Ferguson J. F., Fisher E., McMechan G. A., Sasagawa G., Stevenson J. M., Backus G., Chave A. D., Greer J., Hammer P., Hansen B. L., Hildebrand J. A., Kelty J. R., Sidles C., Wirt J. Test of Newton's Inverse-Square Law in the Greenland Ice Cap, *Physical Review Letters*, 1989, v. 62, 985–988.
14. Thomas J., Vogel P. Testing the Inverse-Square Law of Gravity in Boreholes at the Nevada Test Site, *Physical Review Letters*, 1990, v. 65, 1173–1176.
15. Cahill R. T. Dynamical 3-Space Predicts Hotter early Universe: Resolves CMB-BBN ^{7}Li and ^{4}He Abundance Anomalies, *Progress in Physics*, 2010, v. 1, 67–71.
16. Hertz H. On the Fundamental Equations of Electro-Magnetics for Bodies in Motion, *Wiedemann's Annalen*, 1962, v. 41, 369. Electric Waves, Collection of Scientific Papers, Dover Pub., New York, 1890.
17. Zumberge M. A., Hildebrand J. A., Stevenson J. M., Parker R. L., Chave A. D., Ander M. E., Spiess F. N. Submarine Measurement of the Newtonian Gravitational Constant, *Physical Review Letters*, 1991, v. 67, 3051–3054.

Black Holes in the Framework of the Metric Tensor Exterior to the Sun and Planets

Chifu E. Ndikilar
Department of Physics, Gombe State University, P.M.B 127, Gombe, Gombe State, Nigeria
E-mail: ebenechifu@yahoo.com

The conditions for the Sun and oblate spheroidal planets in the solar system to reduce to black holes is investigated. The metric tensor exterior to oblate spheroidal masses indicates that for the Sun to reduce to a black hole, its mass must condense by a factor of 2.32250×10^5. Using Schwarzschild's metric, this factor is obtained as 2.3649×10^5. Similar results are obtained for oblate spheroidal planets in the solar system.

1 Introduction

It is well known that whenever an object becomes sufficiently compact, general relativity predicts the formation of a black hole: a region of space from which nothing, not even light can escape. The collapse of any mass to the Schwarzschild radius appears to an outside observer to take an infinite time and the events at distances beyond this radius are unobservable from outside, thus the name black hole. From an astronomical point of view, the most important property of compact objects such as black holes is that they provide a superbly efficient mechanism for converting gravitational energy into radiation [1].

The world line element in Schwarzschild's field is well known to be given by [1]

$$c^2d\tau^2 = c^2\left[1-\frac{2GM}{c^2r}\right]dt^2-\left[1-\frac{2GM}{c^2r}\right]^{-1}dr^2 - r^2d\theta^2 - r^2\sin^2\theta d\phi^2. \quad (1)$$

This metric has a singularity, (denoted by r_s) called the Schwarzschild singularity (or radius) at

$$r_s = \frac{2GM}{c^2}. \quad (2)$$

For most physical bodies in the universe, the Schwarzschild radius is much smaller than the radius of their surfaces. Hence for most bodies, there does not exist a Schwarzschild singularity. It is however, speculated that there exist some bodies in the universe with the Schwarzschild radius in the exterior region. Such bodies are called black holes [1].

In this article, the factor by which the radius of the Sun and oblate spheroidal planets is reduced to form a black hole is computed using the oblate spheroidal space-time metric. The results are compared to those obtained using Schwarzschild's metric.

2 Oblate Spheroidal Space-Time Metric

It has been established [2] that the covariant metric tensor in the region exterior to a static homogeneous oblate spheroid in oblate spheroidal coordinates is given as

$$g_{00} = \left(1+\frac{2}{c^2}f(\eta,\xi)\right) \quad (3)$$

$$g_{11} = -\frac{a^2}{1+\xi^2-\eta^2}\left[\eta^2\left(1+\frac{2}{c^2}f(\eta,\xi)\right)^{-1}+\frac{\xi^2(1+\xi^2)}{(1-\eta^2)}\right] \quad (4)$$

$$g_{12}\equiv g_{21} = -\frac{a^2\eta\xi}{1+\xi^2-\eta^2}\left[1-\left(1+\frac{2}{c^2}f(\eta,\xi)\right)^{-1}\right] \quad (5)$$

$$g_{22} = -\frac{a^2}{1+\xi^2-\eta^2}\left[\xi^2\left(1+\frac{2}{c^2}f(\eta,\xi)\right)^{-1}+\frac{\eta^2(1-\eta^2)}{(1+\xi^2)}\right] \quad (6)$$

$$g_{33} = -a^2(1+\xi^2)(1-\eta^2) \quad (7)$$

$$g_{\mu\nu} = 0;\ \text{otherwise}. \quad (8)$$

Thus, the world line element in this field can be written as

$$c^2d\tau^2 = c^2g_{00}dt^2 - g_{11}d\eta^2 - 2g_{12}d\eta d\xi - g_{22}d\xi^2 - g_{33}d\phi^2. \quad (9)$$

Multiplying equation (9) all through by $\left(\frac{1}{dt}\right)^2$ yields

$$c^2\left(\frac{d\tau}{dt}\right)^2 = c^2g_{00} - g_{11}\left(\frac{d\eta}{dt}\right)^2 - 2g_{12}\frac{d\eta}{dt}\frac{d\xi}{dt} - g_{22}\left(\frac{d\xi}{dt}\right)^2 - g_{33}\left(\frac{d\phi}{dt}\right)^2. \quad (10)$$

It can be concluded that the space velocity(v_s) is given as

$$v_s = g_{11}\left(\frac{d\eta}{dt}\right)^2 + 2g_{12}\frac{d\eta}{dt}\frac{d\xi}{dt} + g_{22}\left(\frac{d\xi}{dt}\right)^2 + g_{33}\left(\frac{d\phi}{dt}\right)^2, \quad (11)$$

and the velocity of local time

$$v_\tau = c\frac{d\tau}{dt}. \quad (12)$$

The gravitational velocity can equally be defined with the aid of equation (3) as

$$v_G = \sqrt{-2f(\eta,\xi)}. \tag{13}$$

This implies that

$$c^2 = v_\tau^2 + v_G^2 + v_s^2 \tag{14}$$

or

$$c = \left|\vec{v}_\tau^2 + \vec{v}_G^2 + \vec{v}_s^2\right|. \tag{15}$$

3 Black holes in oblate spheroidal space time of Sun and planets

In the absence of gravity and acceleration, $f(\eta,\xi) = 0$ and thus $v_G = 0$. Hence, v_s can be written explicitly as

$$\begin{aligned} v_s^2 = &\left[\frac{a^2\eta^2}{1+\xi^2-\eta^2} + \frac{\xi^2\left(1+\xi^2\right)}{1-\eta^2}\right]\left(\frac{d\eta}{dt}\right)^2 \\ &+\frac{a^2}{1+\xi^2-\eta^2}\left[\xi^2 + \frac{\eta^2(1-\eta^2)}{1+\xi^2}\right]\left(\frac{d\xi}{dt}\right)^2 \\ &+a^2\left(1+\xi^2\right)\left(1-\eta^2\right)\left(\frac{d\phi}{dt}\right)^2. \end{aligned} \tag{16}$$

Thus in the absence of gravity, equation(14) reduces to

$$\frac{d\tau}{dt} = \sqrt{1-\frac{v_s^2}{c^2}}. \tag{17}$$

It can basically be seen that equation(15) establishes a maximum value of c and hence the gravitational velocity v_G can never exceed c. An approximate expression for $f(\eta,\xi)$ along the equator of an oblate spheroid [3] is

$$f(\eta,\xi) \approx \frac{B_0}{3\xi^2}\left(1+3\xi^2\right)i + \frac{B_2}{30\xi^3}\left(7+15\xi^2\right)i \tag{18}$$

where B_0 and B_2 are constants. Equation(18) can be written equally as

$$f(\eta,\xi) \approx -\left(\frac{C}{\xi} + \frac{D}{\xi^3}\right) \tag{19}$$

where C and D are equally constants. These constants can easily be computed for the oblate spheroidal astrophysical bodies in the solar system and results are presented in Table 1.

Setting the gravitational velocity v_G to be equal to the maximum value c, in equation (13), an approximate expression for the parameter ξ for a black hole in oblate spheroidal space time can be obtained as

$$\xi_{blackhole} \approx \frac{2C}{c^2}. \tag{20}$$

Table 1: Basic constants for oblate spheroidal bodies in the solar system

Body	C [$\times 10^{-9}$ Nmkg^{-1}]	D [$\times 10^{-9}$ Nmkg^{-1}]
Sun	−46796.04	−15598.70
Earth	−0.743851	−0.247962
Mars	−0.1132	−0.03780
Jupiter	−3.77107	−1.25803
Saturn	−0.879543	−0.29356
Uranus	−0.842748	−0.28102
Neptune	−1.065429	−0.35516

Table 2: Reduction ratio for oblate spheroidal masses in the solar system to reduce to black holes

Body	$\xi_{surface}$	$\xi_{blackhole}$	reduction ratio: $\frac{\xi_{surface}}{\xi_{blackhole}}$
Sun	241.52	1.1×10^{-3}	2.32250×10^5
Earth	12.01	1.6×10^{-8}	7.50625×10^8
Mars	09.17	2.0×10^{-9}	4.58500×10^9
Jupiter	02.64	8.3×10^{-8}	3.18070×10^7
Saturn	01.97	1.9×10^{-8}	1.03684×10^8
Uranus	03.99	1.8×10^{-8}	2.21667×10^8
Neptune	04.30	2.3×10^{-8}	1.86950×10^8

Hence the parameter $\xi_{blackhole}$ for various bodies in the solar system is computed using equation (20) and the reduction ratio for oblate spheroidal masses in the solar system to reduce to black holes is obtained (Table 2).

The reduction ratio can equally be calculated using Schwarzschild's expression. The equatorial radius (r) for the bodies is divided by the Schwarzschild's radius (r_{schw}) to obtain the reduction ratio. The results are shown in Table 3.

4 Conclusion

This short article presents the notion of black holes in the metric tensor exterior to oblate spheroidal masses. Equation (20) is an approximate expression for the parameter ξ of an oblate spheroid to collapse to a black hole.Reductions ratios computed using the oblate spheroidal metric for Sun and planets in the Solar system authenticates the soundness of metric. The closeness of the reduction ratio for oblate spheroidal masses in the solar system computed using the metric tensor in oblate spheroidal space time to that in Schwarzschild's metric is remarkable.Basically, since the Sun and planets under consideration are oblate spheroidal in nature, the values obtained using the metric tensor contain slight corrections to values obtained using Schwarzschild's metric.

Submitted on March 21, 2011 / Accepted on March 24, 2011

Table 3: Schwarzschild's reduction ratio

Body	r [$\times 10^3$ m]	r_{schw} [m]	reduction ratio: $\frac{r}{r_{schw}}$
Sun	$700,000$	2.96×10^3	2.36490×10^5
Earth	6378	8.80×10^{-3}	7.24773×10^8
Mars	3396	9.9×10^{-4}	3.43030×10^9
Jupiter	$71,490$	2.8	2.55320×10^7
Saturn	$60,270$	8.5×10^{-1}	7.09059×10^7
Uranus	$25,560$	1.3×10^{-1}	1.96615×10^8
Neptune	$24,760$	1.5×10^{-1}	1.65067×10^8

References

1. Weinberg S. Gravitation and cosmology. J. Wiley, New York, 1972, pp. 175–188.
2. Chifu E. N., Usman A., Meludu O. C. Orbits in Homogeneous Oblate Spheroidal Gravitational Space Time, *Progress in Physics*, 2009, v. 3, 49–53.
3. Chifu E. N., Usman A., Meludu O. C. Gravitational Spectral Shift Exterior to the Sun, Earth and the Other Oblate Spheroidal Planets, *Progress in Physics*, 2010, v. 4, 56–60.

Isotopes and the Electron Configuration of the Blocks in the Periodic Table of Elements, upto the Last Element No.155

Albert Khazan

E-mail: albkhazan@gmail.com

This is a theoretical study, which first manifests which connexion exists between isotopes and the electron blocks, and how the electron blocks are located in the version of the Periodic Table of Elements which ends with element No.155.

1 Introduction

It is known that elements of the Periodic Table of Elements have fractional numerical values of atomic masses. This is because the elements consists of, as regularly, a mix of inborn (native) isotopes. For this reason we conclude that the average weighted atomic mass of all stable isotopes of any element (taking their distribution in the Earth crust) is that atomic mass which is used in all calculations. Because it is equal to the sum of the electric charge of an atomic nucleus and the number of neutrons in it, the isotopes are determined by the condition $A = Z + N$, where A is the atomic mass, Z is the charge, N is the number of neutrons of the nucleus. With all these, it is necessary to keep in mind that, having the same number of protons in a nucleus, the nucleus may contain different number of neutrons which do not change the chemical properties of the atoms: all isotopes of the same element bear the same electric charge of its nucleus, but change only with the number of neutrons in it.

2 Calculation according to the table of isotopes

According to the data provided by *Nuclear Periodic Table* [1], all spectacularity of the data was split into blocks, wherein the number of isotopes was determined, namely: 431 (s), 1277 (p), 1612 (d), 1147 (f). As is seen in Fig. 1, the obtained results form a smooth arc with $R^2 = 1$. Because all the isotopes are grouped into clocks of the electron configuration alike elements of the Periodic Table, we are lawful to conclude that the *same law* lies in the ground of the geometric configurations. It is necessary to note that, with reaching the top of the arc, the number of the isotopes very lowers, that was as well observed in the case of elements of the Periodic Table [2].

3 Version of the Periodic Table of Elements, which limits by element No.155

It is known that the "blocks" of the Periodic Table of Elements are sets of adjacent groups [3, 4]. The names of the blocks originate in the number of the spectroscopic lines of the atomic orbitales in each of them: **s**harp, **p**rincipal, **d**iffuse, **f**undamental. During the last decades, one suggested to extend the Periodic Table upto 218 elements, with appearance a **g**-block in it [5]. If, in the version of the Periodic Table consisting of only 118 elements, the blocks draw a smooth arc with $R^2 = 1$ (see Fig. 2), the appearance of additional elements in the Table requires new construction of the blocks, which should be set up in another configuration.

Earleir [5], we suggested a version of the Periodic Table which contained Period 8 with 37 elements (two lines with 18 and 1 elements in Group 1). In this form, the Periodic Table satisfies the common structure of the location of the elements. However, once lanthanides and actinides have been extended into a common scheme, the heaviest element No.155 (which ends the Table in this its version) became shifted for 4 positions to right. Therefore, a question rose: how to locate these 37 elements in the new version of the Table so that they would completely satisfy all the rules of the electron configuration of the blocks?

First, we added 2 elements to block **s** upto the begining of Period 8. Then we added 6, 10, and 14 elements (respectively) to blocks **p**, **d**, **f**. Concerning the rest 4 elements, we created a new block **g**. All these changes are shown in Fig. 2 (the upper arc). As is seen, the arc has the same form as the lower arc, and shows that fact that the number of elements of the last block reaches the actual limit.

On the basis of that has been said above, a long-period form of the Periodic Table of Elements was constructed by the Author (see Fig. 3). It differs from the hypothetical forms of the Periodic Table by the real data consisting our Table. Element No.155 is the last (heaviest) in our version of the Table, thus this element "closes" the Table. Element No.155 also opens and closes Period 9, being located in Group 1 of this Period.

This scheme of calculation is applicable to all Tables of Elements containing more than 118 elements. The necessity of our study, presented herein, and the suggested version of the Table which limits by element No.155, is due to that fact the law of hyperbolas we used previously in the Periodic Table [5] provided not only the possibility to calculate the upper limit of the Table (element No.155 and its parameters such as atomic mass 411.66), but also allowed to determine its location in the extended version of the Table of Elements.

If earlier the theoretical physical chemists discussed the possibility to add a number of elements over 118 to the Table of Elements (they suggested to do it as new blocks they referred to as superactinide series, eka-superactinide, Ubb-series, Usb-series), we now obviously see that this step is non-

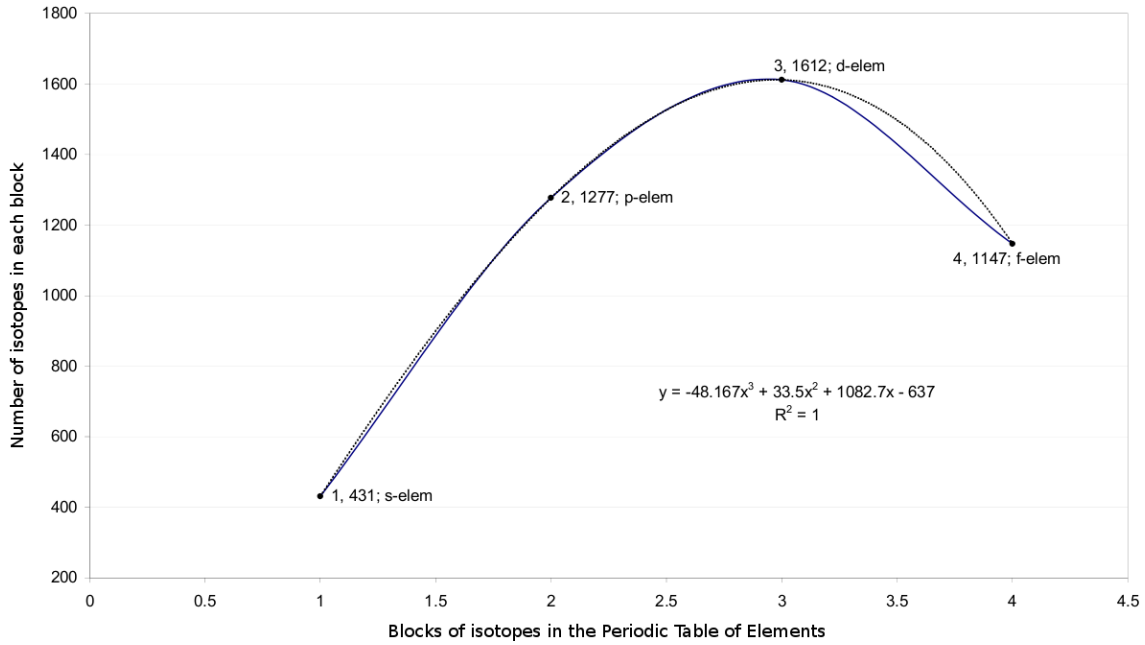


Fig. 1: Dependency of the number of the isotopes in the blocks from their names according to the elements of the electron configuration.

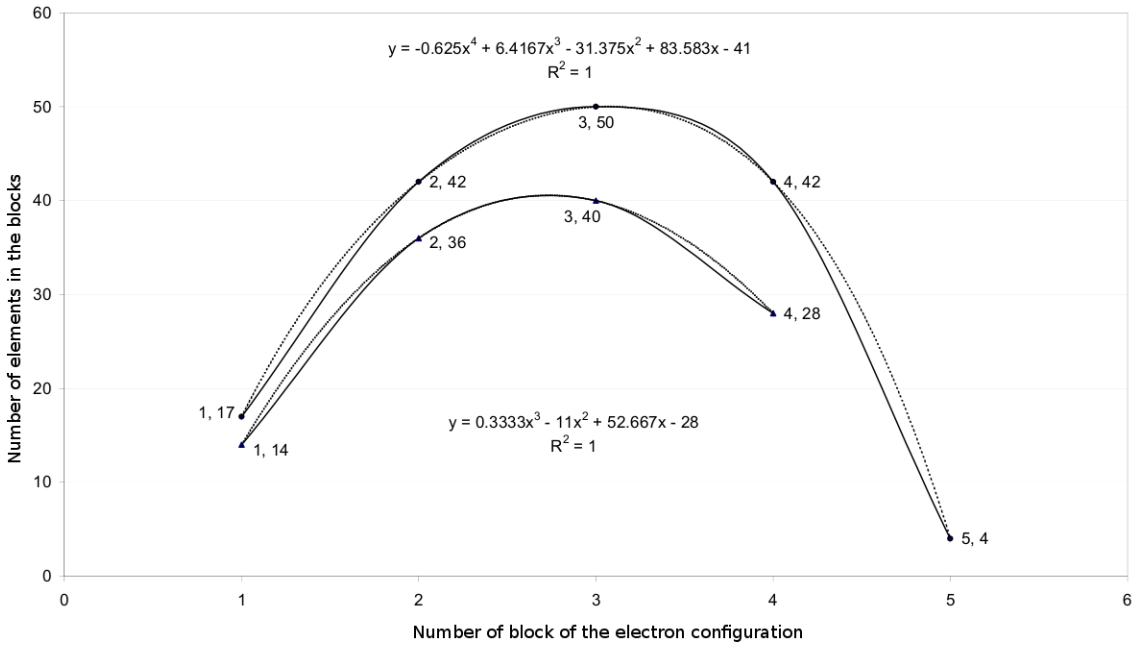


Fig. 2: Results of calculation of the electron configuration of the elements. The lower arc has been calculated for the version of the Periodic Table containing 118 elements. The upper arc has been calculated for the version of the Periodic Table containing 155 elements (suggested by the Author [5]).

1	2																																								
3	4																																						5	6	7	8	9	10
11	12																																					13	14	15	16	17	18
19	20																																					31	32	33	34	35	36
37	38																																					49	50	52	52	53	54
55	56						57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86						
87	88						89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118						
119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154						
155																																										

s-block	g-block	f-block	d-block	p-block

Fig. 3: Periodic Table of Elements, which is limited by element No.155 (suggested by the Author).

sense. Despite the bulky mathematical apparatus of Quantum Mechanics was applied to calculation of stability of the elements, it never led to a result about a limit of the Periodic Table of Elements. This was never claimed in the basis of the quantum mechanical calculations. This is because that the conditions of micro-scales, where the laws of Quantum Mechanics work, do not provide the necessary data for the calculation. Only common consideration of the conditions of micro-world and macro-world, as the author did in the recent study [5], allowed to develop the fundamental law of hyperbolas in the Periodic Table of Elements, which starts from the positions of macro-scale then continues upto the electron configuration of the elements (wherein it works properly as well, as we seen in this paper) that led to that final version of the Periodic Table of Elements, which has been presented in this paper.

Submitted on February 29, 2011 / Accepted on April 04, 2011

References

1. Nuclear Periodic Table from the Radiochemistry Society, http://www.radiochemistry.org/periodictable/NuclearPeriodicTable.html
2. Khazan A. Electron configuration and element No.155 of the Periodic Table of Elements. *Progress in Physics*, 2011, v.2, 39–43.
3. Janet C. La Classification Hélicoïdale des Éléments Chimiques. Beauvais: Imprimerie Départementale de l'Oise, 1928.
4. Casson L. Notice biographique sur la vie et de l'oeuvre de Charles Janet. *Bull. Soc. Acad. de l'Oise*, Beauvais, 2008, 232.
5. Khazan A. Upper Limit in Mendeleev's Periodic Table — Element No.155. Svenska fysikarkivet, Stockholm, 2010.

On the Cold Big Bang Cosmology

Armando V.D.B. Assis

Departmento de Física, Universidade Federal de Santa Catarina — UFSC, Trindade 88040-900, Florianópolis, SC, Brazil.
E-mail: armando.assis@pgfsc.ufsc.br

We solve the general relativity (GR) field equations under the cosmological scope via one extra postulate. The plausibility of the postulate resides within the Heisenberg indeterminacy principle, being heuristically analysed throughout the appendix. Under this approach, a negative energy density may provide the positive energy content of the universe via fluctuation, since the question of conservation of energy in cosmology is weakened, supported by the known lack of scope of the Noether's theorem in cosmology. The initial condition of the primordial universe turns out to have a natural cutoff such that the temperature of the cosmological substratum converges to the absolute zero, instead of the established divergence at the very beginning. The adopted postulate provides an explanation for the cosmological dark energy open question. The solution agrees with cosmological observations, including a 2.7K CMBT prediction.

1 Revisiting the Theoretical Assumptions

The study of the dynamics of the entire universe is known as Cosmology [1–3]. The inherent simplicity in the mathematical treatment of the Cosmology, although the entire universe must be under analysis, should be recognized as being due to Copernicus. Indeed, since the primordial idea permeating the principle upon which the simplicity arises is just an extension of the copernican revolution*: the cosmological principle. This extension, the cosmological principle, just asseverates we are not in any sense at a privileged position in our universe, implying that the average large enough scale† *spatial* properties of the physical universe are the same from point to point at a given cosmological instant. Putting these in a mathematical jargon, one says that the large enough scale spatial geometry at a given cosmological instant t is exactly the same in spite of the position of the observer at some point belonging to this t-sliced three-dimensional universe or, equivalently, that the spatial part of the line element of the entire universe is the same for all observers. Hence, the simplicity referred above arises from the very two principal aspects logically encrusted in the manner one states the cosmological principle:

- The lack of a privileged physical description of the universe at a t-sliced large enough scale $\Rightarrow$ large enough scale $\Rightarrow$ one neglects all kind of known physical interactions that are unimportant on the large enough scales $\Rightarrow$ remains gravity;
- The lack of a privileged physical description of the universe at a t-sliced large enough scale $\Rightarrow$ large enough scale $\Rightarrow$ one neglects local irregularities of a global t-sliced substratum representing the t-sliced universe $\forall$ cosmological instants $t \Rightarrow$ substratum modelled as a fluid without t-sliced spatially localized irregularities $\Rightarrow$ homogeneous and isotropic t-sliced‡ fluid.

One shall verify the t-local characteristic of the the cosmological principle, i.e., that non-privileged description does not necessarily hold on the global time evolution of that t-sliced spacelike hypersurfaces. In other words, two of such t-sliced hypersurfaces at different instants would not preserve the same aspect, as experimentally asseverated by the expansion of the universe. Hence, some further assumption must be made regarding the time evolution of the points belonging to the t-sliced spacelike hypersurfaces:

- The particles of the cosmological fluid are encrusted in spacetime on a congruence of timelike geodesics from a point in the past, i.e., the substratum is modelled as a perfect fluid.

Hence, the following theoretical ingredients are available regarding the above way in which one mathematically construct a cosmological model:

Gravity modeled by Einstein's General Relativity field equations (in natural units):

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}. \tag{1}$$

Homogeneity is mathematically translated by means of a geometry (metric) that is the same from point to point, spatially speaking. Isotropy is mathematically translated by means of a lack of privileged directions, also spatially speaking. These two characteristics easily allow one to consider spaces equipped with constant curvature K. From a differential geometry theorem, Schur's, a n-dimensional space $\mathbb{R}^n$, $n \geq 3$, in which a η-neighbourhood has isotropy $\forall$ points

*Copernicus told us that the Earth is not the center of our planetary system, namely the solar system, pushing down the historical button leading to the collapse of the established *anthropocentric status quo*.

†One must understand large enough scale as being that of cluster of galaxies.

‡One shall rigorously attempt to the fact: the isotropy and homogeneity are t-sliced referred, i.e., these two properties logically emerging from the cosmological principle hold upon the entire fluid at t, holding spatially at t, i.e., homogeneity and isotropy are spatial properties of the fluid. Regarding the time, one observer can be at an own proper τ-geodesic...

belonging to it, has constant curvature K throughout η. Since we are considering, spatially, global isotropy, then K is constant everywhere. Hence, one defines the Riemann tensor:

$$R_{abcd} = K\,(g_{ac}g_{bd} - g_{ad}g_{bc})\,, \tag{2}$$

spatially speaking.

As indicated before, homogeneity and isotropy are spatial properties of the geometry. Time evolution, e.g.: expansion, can be conformally agreed with these two spatial properties logically emerging from the cosmological principle in terms of Gaussian normal coordinates. Mathematically, the spacetime cosmological metric has the form:

$$ds^2 = dt^2 - [a(t)]^2\, d\sigma^2. \tag{3}$$

Since spatial coordinates for a spatially fixed observer do not change, $ds^2 = dt^2 \Rightarrow g_{tt} = 1$.

Regarding the spatial part of the line element, the Schwarzschild metric is spherically symmetric, a guide to our purposes. From the Scharzschild metric (signature $+ - - -$):

$$ds^2 = e^{2\nu(r)}dt^2 - e^{2\lambda(r)}dr^2 - r^2 d\theta^2 - r^2\sin^2\theta\, d\phi^2, \tag{4}$$

one easily writes down the spatial part of the spacetime cosmological metric:

$$d\sigma^2 = e^{2f(r)}dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2. \tag{5}$$

One straightforwardly goes through the tedious calculation of the Christoffel symbols and the components of the Ricci tensor, finding:

$$e^{2f(r)} = \frac{1}{1 - Kr^2}. \tag{6}$$

Absorbing constants* by the scale factor in eqn. (3), one normalizes the curvature constant K, namely $k \in \{-1;\, 0;\, +1\}$. Hence, the cosmological spacetime metric turns out to be in the canonical form:

$$ds^2 = dt^2 - [a(t)]^2\left(\frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2\right). \tag{7}$$

Now, regarding the fluid substratum, one sets in co-moving coordinates ($dt/d\tau = 1$, $u^\mu = (1;\, 0;\, 0;\, 0)$):

$$T^\mu{}_\nu = 0,\ \mu \neq \nu;\ \ T^0{}_0 = \rho;\ \ T^\mu{}_\mu = -p\,,\ \ \text{for } \mu \in \{1;\, 2;\, 3\} \tag{8}$$

since the particles in the fluid are clusters of galaxies falling together with small averaged relative velocities compared with the cosmological dynamics, where the substratum turns out to be averaged described by an average substratum density ρ and by an average substratum pressure p.

The Einstein tensor in eqn. (1), $G_{\mu\nu}$, is related to the Ricci tensor $R_{\mu\nu} = R^\gamma{}_{\mu\gamma\nu}$ (the metric contraction of the curvature tensor (Riemann tensor)), to the Ricci scalar $R = R^\mu{}_\mu$ (the metric contraction of the Ricci tensor) and to the metric $g_{\mu\nu}$ itself:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\,R\,g_{\mu\nu}\,. \tag{9}$$

The curvature tensor $R^\alpha{}_{\beta\gamma\delta}$ is obtained via a metric connection, the Christoffel $\Gamma^\alpha{}_{\beta\delta}$ symbols in our case of non-torsional manifold:

$$R^\alpha{}_{\beta\gamma\delta} = \partial_\gamma\Gamma^\alpha{}_{\beta\delta} - \partial_\delta\Gamma^\alpha{}_{\beta\gamma} + \Gamma^\epsilon{}_{\beta\delta}\Gamma^\alpha{}_{\epsilon\gamma} - \Gamma^\epsilon{}_{\beta\gamma}\Gamma^\alpha{}_{\epsilon\delta}\,, \tag{10}$$

where the metric connection is obtained, in the present case, from the Robertson-Walker cosmological spacetime geometry given by eqn. (7) (from which one straightforwardly obtains the metric coefficients of the diagonal metric tensor in the desired covariant or contravariant representations) via:

$$\Gamma^\alpha{}_{\beta\gamma} = g^{\alpha\delta}\Gamma_{\delta\beta\gamma}\,, \tag{11}$$

being the metric connection (Christoffel symbols) of the first kind $\Gamma_{\delta\beta\gamma}$ given by:

$$\Gamma_{\delta\beta\gamma} = \frac{1}{2}\left(\frac{\partial g_{\beta\gamma}}{\partial x^\delta} + \frac{\partial g_{\gamma\delta}}{\partial x^\beta} - \frac{\partial g_{\delta\beta}}{\partial x^\gamma}\right). \tag{12}$$

These set of assumptions under such mathematical apparatus lead one to the tedious, but straightforward, derivation, via eqn. (1), of the ordinary differential cosmological equations emerging from the relation between the Einstein's tensor, $G_{\mu\nu}$, the Robertson-Walker spacetime cosmological metric of the present case, $g_{\mu\nu}$ via eqn. (7), and the stress-energy tensor, $T_{\mu\nu}$ via metric contraction of the eqn. (8) (signature $+ \ - \ - \ -$):

$$\frac{\dot{R}^2 + kc^2}{R^2} = \frac{8\pi G}{3c^2}\,(\rho + \tilde{\rho})\,; \tag{13}$$

$$\frac{2R\ddot{R} + \dot{R}^2 + kc^2}{R^2} = -\frac{8\pi G}{c^2}\,(p + \tilde{p})\,, \tag{14}$$

where we are incorporating the cosmological constant Λ through the energy density and the pressure of the vacuum: $\tilde{\rho}$ and $\tilde{p}$, respectively. One also must infer we are no more working with natural units. The scale factor becomes $R(t)$, and one must interpret it as the magnification length scale of the cosmological dynamics, since $R(t)$ turns out to be length. This measures how an unitary length of the pervading cosmological substratum at t_0 becomes stretched as the universe goes through a time evolution from t_0 to t. One should not literally interpret it as an increase of the distance between two points, e.g., in a case of expansion, a stretched stationary wavelength connecting two cosmological points at a t_0-sliced spacelike substratum would remain stationarily connecting the very same two points after the stretched evolution to the respective t-sliced spacelike substratum, but less energetically.

*Defining $r' = \sqrt{|K|}\,r$, one straightforwardly goes through...

2 A Cold Beginning?

Applying the following conservation criteria:

$$\nabla_\mu T^\mu_{\ t} = \partial_\mu T^\mu_{\ t} + \Gamma^\mu_{\mu\nu} T^\nu_{\ t} - \Gamma^\nu_{\mu t} T^\mu_{\ \nu} = 0, \tag{15}$$

one finds via the diagonal stress-energy tensor (see eqn. (8)), the metric connection (see eqs. (11) and (12)) and the spacetime cosmological geometry of the present case (eqn. (7)):

$$\frac{\partial}{\partial t}(\rho + \tilde{\rho}) + 3\frac{\dot{R}}{R}(\rho + \tilde{\rho} + p + \tilde{p}) = 0. \tag{16}$$

eqn. (16) is the first law of thermodynamics applied to our substratum (including vacuum), since, despite of geometry, a spatial slice of the substratum has volume $\alpha(k)\,[R(t)]^3$, density $(\rho(t) + \tilde{\rho})^*$ and energy $(\rho(t) + \tilde{\rho})\,\alpha(k)\,[R(t)]^3$, implying that $dE + pdV = 0$ turns out to be eqn. (16). $\alpha(k)$ is the constant that depends on geometry (open, $k = -1$; flat, $k = 0$; closed, $k = 1$) to give the correct volume expression of the mentioned spatial slice of the t-sliced cosmological substratum.

Now, we go further, considering the early universe as being dominated by radiation. In the ultrarelativistic limit, the equation of state is given by:

$$\rho - 3p = 0. \tag{17}$$

Putting this equation of state in eqn. (16) and integrating, one obtains the substratum pressure as a function of the magnification scale R:

$$4\ln\|R\| + \ln\|p\| = C' \;\Rightarrow\; \|p\| = \frac{e^{C'}}{R^4} \;\Rightarrow\; p = \pm\frac{C^+}{R^4}, \tag{18}$$

where $C^+ \geq 0$ is a constant of integration. In virtue of eqn. (18), eqn. (14) is rewritten in a total differential form:

$$2R\dot{R}d\dot{R} + \left(\dot{R}^2 + kc^2 \pm \frac{8\pi G}{c^2}\frac{C^+}{R^2} + \frac{8\pi G}{c^2}\tilde{p}R^2\right)dR = 0. \tag{19}$$

Indeed, eqn. (19) is a total differential of a constant $\lambda(R, \dot{R}) = constant$:

$$d\lambda(R,\dot{R}) = \frac{\partial\lambda(R,\dot{R})}{\partial\dot{R}}d\dot{R} + \frac{\partial\lambda(R,\dot{R})}{\partial R}dR = 0, \tag{20}$$

since:

$$\frac{\partial\lambda(R,\dot{R})}{\partial\dot{R}} = 2R\dot{R} \;\Rightarrow\; \frac{\partial^2\lambda(R,\dot{R})}{\partial R\,\partial\dot{R}} = 2\dot{R}; \tag{21}$$

$$\frac{\partial\lambda(R,\dot{R})}{\partial R} = \dot{R}^2 + kc^2 \pm \frac{8\pi G}{c^2}\frac{C^+}{R^2} + \frac{8\pi G}{c^2}\tilde{p}R^2 \;\Rightarrow \tag{22}$$

$$\frac{\partial^2\lambda(R,\dot{R})}{\partial\dot{R}\,\partial R} = 2\dot{R} \;\therefore\; \frac{\partial^2\lambda(R,\dot{R})}{\partial R\,\partial\dot{R}} = \frac{\partial^2\lambda(R,\dot{R})}{\partial\dot{R}\,\partial R} = 2\dot{R}. \tag{23}$$

Integrating, one has:

$$\int\partial\lambda(R,\dot{R}) = \int 2R\dot{R}\,\partial\dot{R} = 2R\int\dot{R}\,d\dot{R} + h(R) \;\therefore \tag{24}$$

$$\lambda(R,\dot{R}) = R\dot{R}^2 + h(R), \tag{25}$$

where $h(R)$ is a function of R. From eqs. (22) and (25):

$$\frac{\partial}{\partial R}\lambda(R,\dot{R}) = \dot{R}^2 + kc^2 \pm \frac{8\pi G}{c^2}\frac{C^+}{R^2} + \frac{8\pi G}{c^2}\tilde{p}R^2 \Rightarrow$$

$$h(R) = \int\left(kc^2 \pm \frac{8\pi G}{c^2}\frac{C^+}{R^2} + \frac{8\pi G}{c^2}\tilde{p}R^2\right)dR \;\therefore \tag{26}$$

$$h(R) = kc^2R \mp \frac{8\pi G}{c^2}\frac{C^+}{R} + \frac{8\pi G}{3c^2}\tilde{p}R^3. \tag{27}$$

Putting this result from eqn. (27) in eqn. (25):

$$\lambda(R,\dot{R}) = R\dot{R}^2 + kc^2R \mp \frac{8\pi G}{c^2}\frac{C^+}{R} + \frac{8\pi G}{3c^2}\tilde{p}R^3 = constant \tag{28}$$

is the general solution of the total differential equation eqn. (19). Dividing both sides of eqn. (28) by $R^3 \neq 0$:

$$\frac{\lambda(R,\dot{R})}{R^3} = \frac{\dot{R}^2 + kc^2}{R^2} \mp \frac{8\pi G}{c^2}\frac{C^+}{R^4} + \frac{8\pi G}{3c^2}\tilde{p}, \tag{29}$$

using the eqn. (13), one obtains:

$$\frac{\lambda(R,\dot{R})}{R^3} = \frac{8\pi G}{c^2}\left(\frac{\rho}{3} \mp \frac{C^+}{R^4}\right) + \frac{8\pi G}{3c^2}(\tilde{\rho} + \tilde{p}) \;\therefore \tag{30}$$

$$\lambda(R,\dot{R}) = constant = 0, \tag{31}$$

in virtue of eqns. (17), (18) and $\tilde{\rho} + \tilde{p} = 0$ for the background vacuum. Of course, the same result is obtained from eqn. (13), since this equation is a constant of movement of eqn. (14), being eqn. (16) the connection between the two. Neglecting the vacuum contribution in relation to the ultrarelativistic substratum, one turns back to the eqn. (28), set the initial condition $R = R_0$, $\dot{R} = 0$, at $t = 0$, obtaining for the substratum pressure:

$$p(R) = k\frac{c^4R_0^2}{8\pi GR^4}, \tag{32}$$

and for the magnification scale velocity:

$$\dot{R}^2 = -kc^2\left(1 - \frac{R_0^2}{R^2}\right). \tag{33}$$

Now, robustness† requires an open universe with $k = -1$. Hence, the locally flat substratum energy is given by‡:

$$E^+ = -4\pi R^3 p(R) \Rightarrow R_0 = -\frac{2GE_0^+}{kc^4}, \tag{34}$$

*One shall remember the cosmological principle: on average, for large enough scales, at t-sliced substratum, the universe has the same aspect in spite of the spatial localization of the observer in the t-slice $\Rightarrow \rho = \rho(t)$. Also, since Λ is constant, $\tilde{\rho}$ and $\tilde{p}$ are constants such that $\tilde{\rho} + \tilde{p} = 0$.

†For, $\dot{R}^2 \in \mathbb{R}$ in eqn. (33) with $R \geq R_0$.

‡The Hawking-Ellis dominant energy condition giving the positive energy, albeit the expansion dynamics obtained via eqn. (32).

in virtue of eqn. (32) and the initial condition $E^+ = E_0^+$, $R = R_0$ at $t = 0$. Returning to eqn. (33), one obtains the magnification scale velocity:

$$\dot{R} = c\sqrt{1 - \frac{4G^2\left(E_0^+\right)^2}{c^8R^2}}, \tag{35}$$

giving $\dot{R} \to c$ as $R \to \infty$. Rewriting eqn. (35), one obtains the dynamical Schwarzchild horizon:

$$R = \frac{2G}{c^4}\frac{E_0^+}{\sqrt{1 - \dot{R}^2/c^2}}. \tag{36}$$

We will not use the eqn. (34) (now you should read the appendix to follow the following argument) to obtain the energy from the energy density and volume for $t \neq 0$, since we do not handle very well the question of the conservation of energy in cosmology caused by an inherent lack of application of the Noether's theorem. In virtue of the adopted initial conditions, an initial uncertainty R_0 related to the initial spatial position of an arbitrary origin will be translated to a huge uncertainty R at the actual epoch. Indeed, one never knows the truth about the original position of the origin, hence the uncertainty grows as the universe enlarge. The primordial energy from which the actual energy of the universe came from was taken as E_0^+ at the beginning. This amount of energy is to be transformed over the universe evolution, giving the present amount of the universe, i.e., the energy of an actual epoch t-sliced hypersurface of simultaneity. But this energy at each instant t of the cosmological evolution turns out to be the transformed primordial indeterminacy E_0^+, since E_0^+ is to be obtained via the Heisenberg indeterminacy principle. In other words, we argue that the energetic content of the universe at any epoch is given by the inherent indeterminacy caused by the primordial indeterminacy. At any epoch, one may consider a copy of all points pertaining to the same hypersurface of simultaneity but at rest, i.e., an instantaneous non-expanding copy of the expanding instantaneous hypersurface of simultaneity. Related to an actual R indeterminacy of an origin in virtue of its primordial R_0 indeterminacy, one has the possibility of an alternative shifted origin at R. This shifted origin expands with $\dot{R}$ in relation to that non-expanding instantaneous copy of the universe at t. Since the primordial origin was considered to encapsulate the primordial energy E_0^+, this energy at the shifted likely alternative origin should be $E_0^+/\sqrt{1 - \dot{R}^2/c^2}$, since, at R, a point expands with $\dot{R}$ in relation to its non-expanding copy. We postulate:

- The actual energy content of the universe is a consequence of the increasing indeterminacy of the primordial era. Any origin of a co-moving reference frame within the cosmological substratum has an inherent indeterminacy. Hence, the indeterminacy of the energy content of the universe may create the impression that the universe has not enough energy, raising illusions as dark energy and dark matter speculations. In other words, since the original source of energy emerges as an indeterminacy, we postulate this indeterminacy continues being the energy content of the observational universe: $\delta E(t) = E^+(t) = E_0^+/\sqrt{1 - \dot{R}^2/c^2}$.

This result is compatible with the Einstein field equations. The compatibility is discussed within the appendix. In virtue of this interpretation, eqn. (36) has the aspect of the Schwarzchild radius, hence the above designation. The t-instantaneous locally flat spreading out rate of dynamical energy at t-sliced substratum is given by the summation over the ν-photonic frequencies:

$$\dot{R}\frac{d}{dR}\left(\frac{E_0^+}{\sqrt{1 - \dot{R}^2/c^2}}\right) =$$

$$= \frac{8\pi^2R^2h}{c^2}\int_0^\infty \frac{\nu^3}{\exp\left(h\nu/k_BT\right) - 1}d\nu = \frac{8\pi^6k_B^4R^2}{15c^2h^3}T^4, \tag{37}$$

where k_B is the Boltzmann constant, h the Planck constant and T the supposed rapid thermodynamically equilibrated t-sliced locally flat instantaneous cosmological substratum temperature. Now, setting, in virtue of the Heisenberg principle:

$$\frac{E_0^+R_0}{c} \approx h \overset{(34)}{\Rightarrow} \left(E_0^+\right)^2 = \frac{hc^5}{2G}, \tag{38}$$

one obtains, in virtue of eqn. (37):

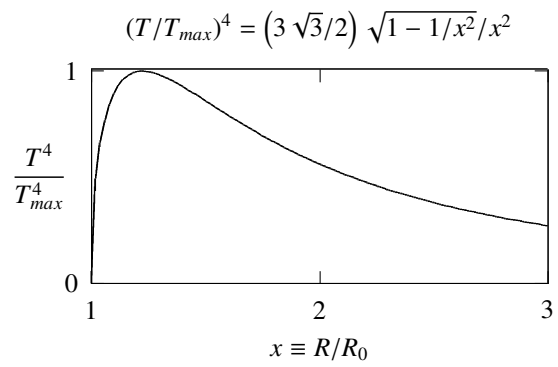
$$T^4 = \frac{15c^7h^3}{16\pi^6Gk_B^4}\frac{1}{R^2}\sqrt{1 - \frac{2Gh}{c^3R^2}}. \tag{39}$$

Hence, the temperature of the cosmological substratum vanishes* at $t = 0$, rapidly reaching the maximum $\approx 10^{32}K$, and assintotically decreasing to zero again as $t \to \infty$.

Indeed. $R_0 = R(t = 0) = \sqrt{2Gh/c^3}$, in virtue of eqs. (34) and (38), giving $T^4(R_0) = T^4(t = 0) = 0$. Also, the maximum temperature is $T \approx 10^{32}K$, from eqn. (39), occuring when $R = R_{max} = \sqrt{3/2}R_0 = \sqrt{3Gh/c^3}$, as one obtains by $dT^4/dR = 0$ with $d^2T^4/dR^2 < 0$. Below†, one infers these properties of eqn. (39).

*We argue there is no violation of the third law of thermodynamics, since one must go from the future to the past when trying to reach the absolute zero, violating the second law of thermodynamics. At $t = 0$, one is not reaching the absolute zero since there is no past before the beginning of the time. To reach the absolute zero, in an attempt to violate the Nernst principle, one must go from the past to the future.

†The eqn. (39) is simply rewritten to plot the graph, i.e.: $T_{max}^4 = \left(5\sqrt{3}\,c^{10}h^2\right)/\left(48\pi^6G^2k_B^4\right)$ and, as obtained before, $R_0 = \sqrt{2Gh/c^3}$.



Now, one puts the result of eqn. (38) in eqn. (35) and integrates:

$$\int_{(2Gh/c^3)^{1/2}}^{R} \frac{R}{\sqrt{R^2 - 2Gh/c^3}} dR = c\int_0^t d\tau, \qquad (40)$$

obtaining:

$$t = \frac{1}{c}\sqrt{R^2 - 2Gh/c^3} \Rightarrow t(R_{max}) = \sqrt{\frac{Gh}{c^5}} \approx 10^{-43}\,\mathrm{s}, \qquad (41)$$

for the elapsed time from $t = 0$ to the instant in which the substratum temperature reaches the maximum value $T \approx 10^{32}$ K. The initial acceleration, namely the explosion/ignition acceleration at $t = 0$ of the substratum is obtained from eqn. (35):

$$\ddot{R} = \dot{R}\frac{d\dot{R}}{dR} = \frac{4G^2\left(E_0^+\right)^2}{c^6R^3} \overset{(38)}{=} \frac{2Gh}{cR^3} \therefore \qquad (42)$$

$$\ddot{R}\left(R = R_0 = \sqrt{2Gh/c^3}\right) = \sqrt{\frac{c^7}{2Gh}} \approx 10^{51}\,\mathrm{m/s^2}. \qquad (43)$$

An interesting calculation is the extension of the eqn. (39) formula to predict the actual temperature of the universe. Since $2Ghc^{-3}R^{-2} \ll 1$ for actual stage of the universe, eqn. (39) is approximately given by:

$$T^4 \approx \frac{15c^7h^3}{16\pi^6Gk_B^4}\frac{1}{R^2} \Rightarrow R^2 \approx \frac{15c^7h^3}{16\pi^6Gk_B^4}\frac{1}{T^4}. \qquad (44)$$

Also, for actual age of the universe, eqn. (41) is approximately given by:

$$t \approx \frac{R}{c} \overset{(44)}{=} \sqrt{\frac{15c^5h^3}{16\pi^6Gk_B^4}}\frac{1}{T^2} \quad \therefore \qquad (45)$$

$$T^2_{\mathrm{Now}} = \sqrt{\frac{15c^5h^3}{16\pi^6Gk_B^4}}\, t^{-1}_{\mathrm{Now}} = 5.32\times10^{20}t^{-1}_{\mathrm{Now}}\left(\mathrm{K^2s}\right). \qquad (46)$$

Before going further on, one must remember we are not in a radiation dominated era. Hence, the left-hand side and the right-hand side of eqn. (37) must be adapted for this situation. The left-hand accomplishes the totality of spreading out energy in virtue of cosmological dynamics. It equals the right-hand side in an ultrarelativistic scenario. But, as the universe evolves, the right-hand side becomes a fraction of the totality of spreading out energy. Rigorously, as the locally flatness of the t-sliced substratum increases, one multiplies both sides of eqn. (37) by $(4/c)\times\left(1/4\pi R^2\right)$ and obtains the t-sliced instantaneously spreading out enclosed energy density. Hence the right-hand side of eqn. (37) turns out to be multiplied by the ratio between the total cosmological density* ρ_c and the radiation density ρ_r. Hence, eqn. (46) is rewritten:

$$\sqrt{\frac{\rho_c}{\rho_r}}\, T^2_{\mathrm{Now}} = 5.32\times10^{20}t^{-1}_{\mathrm{Now}}\left(\mathrm{K^2s}\right). \qquad (47)$$

The actual photonic density is $\rho_r = 4.7\times10^{-31}\,\mathrm{kg/m^3}$ and the actual total cosmological density is $\rho_c = 1.3\times10^{-26}\,\mathrm{kg/m^3}$. For the reciprocal age of universe, t^{-1}_{Now} in eqn. (47), one adopts the Hubble's constant, for open universe, $H = t^{-1}_{\mathrm{Now}} = 2.3\times10^{-18}\,\mathrm{s^{-1}}$. Hence, by eqn. (47), one estimates the actual temperature of the universe:

$$T^2_{\mathrm{Now}} = \sqrt{\frac{4.7\times10^{-31}}{1.3\times10^{-26}}}\times5.32\times10^{20}\times2.3\times10^{-18}\ \mathrm{K^2} \therefore \qquad (48)$$

$$T_{\mathrm{Now}} = 2.7\,\mathrm{K}, \qquad (49)$$

very close to the CMB temperature.

3 Appendix

From eqns. (17) and (32):

$$\rho = 3p = -\frac{3c^4R_0^2}{8\pi G}\frac{1}{R^4} \Rightarrow E_\rho = -\frac{c^4R_0^2}{2G}\frac{1}{R}, \qquad (50)$$

since $k = -1$; E_ρ is the energy (negative) obtained from volume and ρ. From eqn. (34), $R_0^2 = 4G^2(E_0^+)^2/c^8$. Hence, eqn. (50) is rewritten:

$$E_\rho = -\frac{2G}{c^4}\left(E_0^+\right)^2\frac{1}{R}. \qquad (51)$$

With the eqn. (36), we reach:

$$E_\rho = -E_0^+\sqrt{1-\dot{R}^2/c^2}. \qquad (52)$$

This negative energy arises from the adopted negative pressure solution. But, its fluctuation is positive:

$$\delta E_\rho = \frac{E_0^+}{\sqrt{1-\dot{R}^2/c^2}}\frac{\dot{R}\,\delta\dot{R}}{c^2}, \qquad (53)$$

since both, $\dot{R}$ and $\delta\dot{R}$, are positive within our model (see eqn. (40)). Let δt be the time interval within this fluctuation process. Multiplying both sides of the eqn. (53) by δt, we obtain:

$$\delta E_\rho\,\delta t = \frac{E_0^+}{\sqrt{1-\dot{R}^2/c^2}}\left(\dot{R}\delta\dot{R}/c^2\right)\delta t. \qquad (54)$$

*Actually, the critical one, since observations asseverate it.

The above relation must obey the Heisenberg indeterminacy principle, and one may equivalently interpret it under the following format:

$$\delta E_\rho \, \delta t = \frac{E_0^+}{\sqrt{1 - \dot{R}^2/c^2}} \, (\delta t)^* \approx h, \tag{55}$$

An energy indeterminacy having the magnitude of the actual cosmological energy content carries an indeterminacy $\delta\dot{R} \approx c$ about the magnification scale velocity $\dot{R}$ with $\dot{R} \approx c$. For such an actual scenario in which $\dot{R} \approx c$ (see eqn. (35) with $R \to \infty$), we have:

$$\delta t \approx (\delta t)^* \Rightarrow \delta E_\rho \Big|_{R_0}^{\infty} = E^+ = \frac{E_0^+}{\sqrt{1 - \dot{R}^2/c^2}}, \tag{56}$$

if* $\dot{R} \to c$. Now, let's investigate the primordial time domain $t \approx 0$. To see this, we rewrite $\dot{R}\delta\dot{R}$ within the eqn. (54). Firstly, from eqn. (35):

$$\dot{R} = c\sqrt{1 - R_0^2/R^2} \Rightarrow \dot{R}\,\delta\dot{R} = \frac{c^2 R_0^2}{R^3}\,\delta R, \tag{57}$$

where $R_0 = \sqrt{2Gh/c^3}$ as obtained before. Within the primordial time domain $t \approx 0$, we have $R \approx R_0$ and $\delta R \approx R_0$, as discussed before. Hence, the eqn. (57) reads:

$$\dot{R}\,\delta\dot{R} \approx c^2. \tag{58}$$

if $t \approx 0$. Back to the eqn. (54) we obtain again:

$$\delta t \approx (\delta t)^* \Rightarrow \delta E_\rho \Big|_{\approx R_0} = E^+ = \frac{E_0^+}{\sqrt{1 - \dot{R}^2/c^2}}, \tag{59}$$

if $t \approx 0$. This justify the use of $E^+ = E_0^+/\sqrt{1 - \dot{R}^2/c^2}$ within our postulate, emerging from the positive fluctuation of the negative energy E_ρ obtained from volume and the negative energy density ρ stated via the fluid state equation, eqn. (17), and entering within the field equations.

Acknowledgements

A.V.D.B.A is grateful to Y.H.V.H and CNPq for financial support.

Submitted on April 10, 2011 / Accepted on April 13, 2011

References

1. Bondi H. Cosmology. Dover Publications, Inc., New York, 2010.
2. Bondi H. Negative mass in General Relativity. *Review of Modern Physics*, 1957, v. 29 (3), 423–428.
3. Carrol S. Spacetime and Geometry. An Introduction to General Relativity. Addison Wesley, San Francisco, 2004.

*Eqn. (56) holds from $t > 10^{-43}$ seconds, as one easily verify from eqn. (35).

LETTERS TO PROGRESS IN PHYSICS

LETTERS TO PROGRESS IN PHYSICS

Arthur Marshall Stoneham

(1940–2011)

Vahan Minasyan and Valentin Samoilov
Scientific Center of Applied Research, JINR, Joliot-Curie 6, Dubna, 141980, Russia
E-mails: mvahan@scar.jinr.ru; scar@off-serv.jinr.ru

The memory of the prominent British physicist, Prof. Arthur Marshall Stoneham (1940–2011), will live in our hearts and souls.

Arthur Marshall Stoneham (1940–2011)

Marshall Stoneham was born in 1940 in Barrow-in-Furness, Cumbria. He was educated at Barrow Grammar School for Boys before reading physics at Bristol University. In 1964 he completed his doctorate at Bristol under Prof. Maurice Pryce. After completing his thesis, Marshall started working for the Atomic Energy Authority in the Theoretical Physics Division at Harwell. At that time, Harwell faced challenges posed by the nuclear programme, involving the construction of reactors and the safe disposal of radioactive waste. Marshall's main work was the Theory of Defects in Solids. His book on the subject left its mark on a generation, aided by Marshall's habit of referring people to the precise place in the book where their answer was. It has never been out of print since it was first published. Marshall's group at Harwell became a leading light for both the nuclear industry and beyond. He became a division head, AEA chief scientist and retained his interest in nuclear power (both fission and fusion) to the end.

At University College London (where he moved in 1995 as Massey Professor of Physics) he and his colleague John Finney built up the London Centre of Nanotechnology. He was an Honorary Fellow of Wolfson College, Oxford University, from 1985, was elected a Fellow of the Royal Society in 1989 and in 2010 had started his term as President of the Institute of Physics. His colleagues will remember him for his support (even when that support took the form of asking a killer question at the end of your presentation after apparently having slept through it), for the way he promoted their work — even if it was by remarking, "Oh they could sort that out in a few days". It never took less than three months. Above all, in a life that was filled to overflowing he found time for people; to listen, to encourage, to advise. Marshall was a prolific writer. In addition to several books, he was author or co-author of over 500 publications.

Marshall had a great love of music and played the French horn, inspired by a recording of Dennis Brain playing the Mozart horn concertos which his father bought for him (and regretted!) at the age of 18. His love of wind music led him to form his own music group in 1971, the Dorchester Wind Players. Throughout the '80s and '90s he dedicated himself to the massive task of compiling a directory of every piece of wind music in the world ever written for two or more instruments, into a *Wind Ensemble Sourcebook*. Marshall's professional life took him all over the world and he used these travel opportunities to rummage in obscure music libraries and even monasteries in his quest. The project took years, but eventually he and his co-authors published *Wind Ensemble Sourcebook* in 1997. It runs to 450 pages, containing records of 12,000 works by 2,200 composers, and it has two companion volumes. The whole enterprise was truly a world first and will probably never be equalled.

In 1962 Marshall married Doreen, another physicist, and also from Barrow-in-Furness. They have two physicist daughters and he would often joke that he had "done his bit" for women in science. Marshall was a director in his and his wife's specialist laboratory, Oxford Authentication, which authenticates pottery and porcelain antiquities using thermoluminescence dating.

Marshall died on 18 February from complications arising from pancreatic cancer.

Commencing in 2002 Prof. Stoneham helped us by scientific way. He had given scientific directions which then

became the basis of our research studies. Due to his scientific support, we published many important papers in the science connected with understanding structures of light and solid. We never forget Arthur Marshall Stoneham who was a very noble man. His memory will live always in our souls.

We are very grateful to Doreen Stoneham who helped us by information connected with the early private life of Prof. Stoneham.

We also thank the Editor of Chief of Progress in Physics, Dr. Dmitri Rabounski, who assisted us with this letter, and always helped us by scientific way.

Submitted on April 04, 2011 / Accepted on April 09, 2011

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), Zentralblatt MATH (Germany), 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,
200 College Road, Gallup, NM 87301, USA

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

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

JULY 2011 **VOLUME 3**

SPECIAL ISSUE

The Sun — Gaseous or Liquid? A Thermodynamic Analysis

CONTENTS

Information for Authors and Subscribers

Progress in Physics has been created for 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. 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. We accept brief papers, no larger than 8 typeset journal pages. Short articles are preferable. Large papers can be considered in exceptional cases to the section *Special Reports* intended for such publications in the journal. 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.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

A Thermodynamic History of the Solar Constitution — I: The Journey to a Gaseous Sun

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Suite 302, Columbus, Ohio 43210, USA
E-mail: robitaille.1@osu.edu

History has the power to expose the origin and evolution of scientific ideas. How did humanity come to visualize the Sun as a gaseous plasma? Why is its interior thought to contain blackbody radiation? Who were the first people to postulate that the density of the solar body varied greatly with depth? When did mankind first conceive that the solar surface was merely an illusion? What were the foundations of such thoughts? In this regard, a detailed review of the Sun's thermodynamic history provides both a necessary exposition of the circumstance which accompanied the acceptance of the gaseous models and a sound basis for discussing modern solar theories. It also becomes an invitation to reconsider the phase of the photosphere. As such, in this work, the contributions of Pierre Simon Laplace, Alexander Wilson, William Herschel, Hermann von Helmholtz, Herbert Spencer, Richard Christopher Carrington, John Frederick William Herschel, Father Pietro Angelo Secchi, Hervé August Etienne Albans Faye, Edward Frankland, Joseph Norman Lockyer, Warren de la Rue, Balfour Stewart, Benjamin Loewy, and Gustav Robert Kirchhoff, relative to the evolution of modern stellar models, will be discussed. Six great pillars created a gaseous Sun: 1) Laplace's Nebular Hypothesis, 2) Helmholtz' contraction theory of energy production, 3) Andrew's elucidation of critical temperatures, 4) Kirchhoff's formulation of his law of thermal emission, 5) Plücker and Hittorf's discovery of pressure broadening in gases, and 6) the evolution of the stellar equations of state. As these are reviewed, this work will venture to highlight not only the genesis of these revolutionary ideas, but also the forces which drove great men to advance a gaseous Sun.

1 On the history of solar science

Pondering upon the history of solar science [1–14], it becomes apparent that, in every age, the dominant theory of the internal constitution of the Sun reflected the state of human knowledge. As understanding of the physical world grew, the theories of old were slowly transformed. Eventually, under the burden of evidence, ancient ideas were destined to disappear completely from the realm of science, relinquished to the sphere of historical curiosity [2]. What was once considered high thought, became discarded.

If science is to advance, historical analysis must not solely reiterate the progress of civilization. Its true merit lies not in the reminiscence of facts, the restatement of ancient ideas, and the reliving of time. Rather, scientific history's virtue stems from the guidance it can impart to the evolution of modern research.

Historical compilations, dissected with contemporary scientific reasoning, have the power to expose both the truths and the errors which swayed our formation of a gaseous Sun [15–21]. These models have evolved as a direct manifestation of mankind's physical knowledge in the 19th and 20th centuries. Through historical review, it can be demonstrated that virtually every salient fact which endowed the Sun with a gaseous interior has actually been refuted or supplanted by modern science. Astrophysics, perhaps unaware of the historical paths followed by its founders [1–14], has at times overlooked the contributions and criticisms of "non-astronomers". Perhaps unable to accept the consequences stemming from the discoveries of the present age, it has continued to perpetuate ideas which can no longer hold any basis in the physical world.

2 Pillars of a gaseous Sun

Five great pillars gave birth to the gaseous Sun in the middle and late 19th century. They were as follows: 1) Laplace's nebular hypothesis [22, 23], 2) Helmholtz' contraction theory [24, 25], 3) Cagniard de la Tour's discovery of critical phenomena [26, 27] and Andrew's elucidation of critical temperatures [28, 29], 4) Kirchhoff's formulation of his law of thermal emission [30–32], and 5) the discovery of pressure broadening in gases by Plücker, Hittorf, Wüllner, Frankland, and Lockyer [33–37]. Today, the last four of these pillars have collapsed, either as scientifically unsound (pillar 4), or as irrelevant with respect to discussions of the internal constitution of the Sun and the nature of the photosphere (pillars 2, 3, and 5). Only the first argument currently survives as relevant to solar theory, albeit in modified form. Nevertheless, each of these doctrines had acted as a driving force in creating a gaseous Sun. This was especially true with regards to the ideas advanced by Helmholtz, Andrews, Kirchhoff, and those

who discovered pressure broadening.

A careful scrutiny of history reveals that, beyond these factors, the greatest impulse driving mankind to a gaseous Sun was the power of theoretical models. In fact, given that all the great experimental forces have evaporated, astrophysics is left with the wonder of its theoretical formulations. Hence, a 6th pillar is introduced: the stellar equations of state [15–17]. It is an important foundation, one which remains intact and whose influence continues to dominate virtually every aspect of theoretical astrophysics.

2.1 Laplace's nebular hypothesis

Laplace's nebular hypothesis [22,23] was often proposed as a starting point for stellar formation in the 19th century. It became the seed for Helmholtz' contraction theory [24, 25], as will be seen in Section 2.2. Laplace's hypothesis was based on the idea that the Sun and the solar system were created by the slow contraction of a nebulous mass. It was initially outlined in very general terms [38] by Emanuel Swedenborg [39, p. 240–272]. Swedenborg, a Swedish philosopher and theologian, believed himself capable of supernatural communication [40, p. 429]. He made numerous contributions to the natural sciences, but in astronomy, the ideas which brought forth the nebular hypothesis may not be solely his own. Rather, Swedenborg might have simply restated the thoughts of the ancient philosophers [2, 38–40]. Still, for the astronomers of the 19th century, Laplace's name stands largely alone, as the father of the nebular hypothesis.

At present, the Solar Nebular Disk Model (SNDM) [41] has largely replaced the nebular hypothesis, although it maintains, in part, its relationship with the original ideas of Laplace. Space limitation prevents our discussion of these concepts. The point is simply made that, despite the passage of more than two centuries, there remains difficulties with our understanding of the formation of the solar system, as Woolfson recalls: "*In judging cosmogonic theories one must have some guiding principle and that oft-quoted adage of the fourteenth-century English monk, William of Occam, known as Occam's razor, has much to commend it. It states 'Essentia non sunt multiplicanda praeter necessitatem' which loosely translates as 'the simplest available theory to fit the facts is to be preferred'. The characteristics of the SNDM is that it neither fits the facts nor is it simple*" [42].

As for Laplace's nebular hypothesis, it was never specific to a particular solar phase (gas, liquid, or solid). Thus, even Kirchhoff had recourse to the ideas of Laplace in arguing for a solid or liquid photosphere [43, p. 23]. The theory could be applied to all solar models and finds prominence in many discussions of solar formation throughout the 19th century. Logically, however, the concept of a slowly contracting gaseous nebular mass enabled a continuous transition into Helmholtz's theory and the stellar equations of state. This was an aspect not shared by the liquid or solid models of the Sun. Hence, Laplace's ideas, though not counter to the liquid or solid Sun, were more adapted to a gaseous solar mass.

2.2 Helmholtz' contraction theory

Helmholtz' great contraction theory dominated solar science almost since the time it was elucidated at a Königsberg lecture on February 7th, 1858 [24, 25]. The mathematical essence of this lecture was rapidly reprinted in its entirety [24]. Prior to the birth of this theory, solar energy production was based on the meteoric hypothesis as introduced by J.R. Mayer [44], one of the fathers of the 1st law of thermodynamics [45]. The meteoric hypothesis was then championed by Lord Kelvin [46, 47]. Hufbauer provided an excellent description of the evolution of these ideas [14, p. 55–57]. Despite the statures of Mayer [44,45] and Thomson [46,47], the meteoric hypothesis quickly collapsed with the dissemination of Helmholtz' work [24, 25]. The contraction theory became a dominant force in guiding all solar models from the middle of the 19th century through the beginning of the 20th. Given the relative incompressibility of liquids and solids, Helmholtz' concepts were more compatible with the gaseous models. The 1660 law of Boyle [48] and the law of Charles [49], published in 1802 by Gay-Lussac, had just been combined into ideal gas law by Claperon in 1832 [50]. Consequently, it was more logical to assume a gaseous interior. Helmholtz' theory was consequently destined to prominence.

When formulating his contraction hypothesis, Helmholtz emphasized the contraction of nebular material, as advanced by Laplace [24, p. 504]. He stated: "*The general attractive force of all matter must, however, impel these masses to approach each other, and to condense, so that the nebulous sphere became incessantly smaller, by which, according to mechanical laws, a motion of rotation originally slow, and the existence of which must be assumed, would gradually become quicker and quicker. By the centrifugal force, which must act most energetically in the neighborhood of the equator of the nebulous sphere, masses could from time to time be torn away, which afterwards would continue their courses separate from the main mass, forming themselves into single planets, or, similar to the great original sphere, into planets with satellites and rings, until finally the principle mass condensed itself into the Sun*" [24, p. 504–505].

The contraction theory of energy production would not easily yield its pre-eminent position in solar science, surviving well into the 20th century. Still, practical difficulties arose with Helmholz' ideas, particularly with respect to the age of the Earth. Eventually, the concept became outdated. Nuclear processes were hypothesized to fuel the Sun by Arthur Eddington in his famous lecture of August 24th, 1920 [51]. This dramatic change in the explanation of solar energy production [52] would produce no obstacle to maintaining a gaseous Sun. This was true even though Helmholtz' theory had been so vital to the concept of a gaseous interior, both in its inception and continued acceptance. Astrophysics quickly abandoned Helmholtz' contraction hypothesis and adopted an al-

ternative energy source, without any consequence for the internal constitution of the Sun. Ultimately, the advantages of condensed matter in solar fusion were never considered. This remained the case, even though the internuclear proximity within the solid or liquid might have held significant theoretical advantages for fusion when combined with the enormous pressures inside the Sun.

2.3 Andrews and critical temperatures

Addressing the role of Andrews and critical temperatures [28, 29] for solar theory, Agnes Clerke stated: "*A physical basis was afforded for the view that the Sun was fully gaseous by Cagniard de la Tour's experiments of 1822, proving that, under conditions of great heat and pressure, the vaporous state was compatible with considerable density. The position was strengthened when Andrews showed, in 1869, that above a fixed limit of temperature, varying for different bodies, true liquefaction is impossible, even though the pressure be so tremendous as to retain the gas within the same space that enclosed the liquid*" [11, p. 188]. A. J. Meadows echoed these ideas when he later added: "*Andrews showed that there existed a critical temperature for any vapour above which it could not be liquefied by pressure alone. This was accepted as confirming the idea, evolved in the 1860's, of a mainly gaseous Sun whose gas content nevertheless sometimes attained the density and consistency of a liquid*" [13, p. 30].

In the second half of the 19th century, the interior of the Sun was already hypothesized to be at temperatures well exceeding those achievable on Earth in ordinary furnaces. It became inconceivable to think of the solar interior as anything but gaseous. Hence, the gaseous models easily gained acceptance. Even today, it is difficult for some scientists to consider a liquid sun, when confronted with a critical temperature for ordinary hydrogen of $-240.18\,\mathrm{C}$, or $\sim 33\,\mathrm{K}$ [53, p. 4–121]. In view of this fact, the existence of a liquid photosphere seems to defy logic.

However, modern science is beginning to demonstrate that hydrogen can become pressure ionized such that its electrons enter metallic conductions bands, given sufficiently elevated pressures. Liquid metallic hydrogen will possess a new critical temperature well above that of ordinary hydrogen. Already, liquid metallic hydrogen is known to exist in the modern laboratory at temperatures of thousands of Kelvin and pressures of millions of atmospheres [54–56]. The formation of liquid metallic hydrogen brings with it a new candidate for the constitution of the Sun and the stars [57–60]. Its existence shatters the great pillar of the gaseous models of the Sun which the Andrew's critical point for ordinary gases [28, 29] had erected. It seems that the phase diagram for hydrogen is much more complex than mankind could have imagined in the 19th century. The complete story, relative to hydrogen at high temperatures and pressures, may never be known. Nevertheless, it is now certain: the foundation built by Andrews [28] has given way.

2.4 Kirchhoff's law of thermal emission

Gustav Kirchhoff thought that the solar photosphere was either liquid or solid [43]. He based his belief on the continuous nature of the solar spectrum, adding that its generation by condensed matter was "*the most probable proposition*" [43]. In hindsight, Kirchhoff should have been even more forceful, as the existence of a continuous solar spectrum produced by condensed matter was indeed *the only possible proposition*. Kirchhoff held the answer in his hands nearly 150 years ago, but through the erroneous formulation [61–66] of his law of thermal emission [30–32] he allowed his insight on the state of the photosphere to be usurped by scientific error.

In speaking on the physical constitution of the Sun, Kirchhoff referred to his law of thermal emission in stating: "*for all bodies begin to glow at the same temperature. Draper has ascertained experimentally the truth of this law for solid bodies, and I have given a theoretical proof for all bodies which are not perfectly transparent; this, indeed, follows immediately from the theorem, concerning the relation between the power of absorption and the power of emission of all bodies*" [43, p. 26]. Of course, Kirchhoff's extension of Draper's findings from solid bodies to liquids and gases enabled the creation of a fully gaseous Sun in the 20th century. Kirchhoff's law stated that, within an adiabatic or isothermal opaque cavity at thermal equilibrium, the radiation would always be represented by a universal blackbody spectrum whose appearance was solely dependent on temperature and frequency of observation, irrespective of the nature of the walls (provided that they were not transparent) or the objects they contained [30–32]. Kirchhoff's law argued, by extension, that a gas could produce a continuous blackbody spectrum. Provided that the Sun could be conceived as following the restrictions for enclosure as required by Kirchhoff's law, there could be no problems with a gaseous structure for the production of the continuous solar spectrum. As such, Kirchhoff had already condemned his liquid photosphere [43] three years earlier, when he formulated his "law of thermal emission" [30–32]. According to Kirchhoff's law, liquids and solids were not required to obtain a blackbody spectrum. This unintended error would permeate physics throughout the next 150 years.

The problems with Kirchhoff's law were not simple to identify [61–66] and Planck himself [67, 68] echoed Kirchhoff's belief in the universal nature of radiation under conditions of thermal equilibrium [69, p. 1–25]. Planck did not discover Kirchhoff's critical error. Furthermore, his own derivation of Kirchhoff's law introduced arguments which were, unfortunately, unsound (see [61, 64, 65] for a complete treatment of these issues). In reality, the universality promoted by Kirchhoff's law involved a violation of the first law of thermodynmaics, as the author has highlighted [65, p. 6].

The acceptance of Kirchhoff's law, at the expense of Stewart's correct formulation [70], enabled the existence of a gaseous Sun. Its correction [61–66] immediately invalidates

the existence of a gaseous photosphere. Condensed matter is required to produce a continuous thermal spectrum, such as that emitted by the solar photosphere. Blackbody radiation was never universal, as Kirchhoff advocated [30–32] and much of astrophysics currently believes. If Kirchhoff's law had been valid, scientists would not still be seeking to understand the nature of the solar spectrum [71–73] after more than 150 years [74–76]. In reality, the most important pillar in the erection of a gaseous Sun was defective.

2.5 Pressure broadening

Despite the existence of Kirchhoff's law, physicists in the early 1860's understood that gases did not produce continuous spectra. Gases were known to emit in lines or bands. As a result, though Kirchhoff's law opened the door to a gaseous Sun, it was not supported by sound experimental evidence. It was under these circumstances, that the concept of pressure broadening in gases entered astrophysics.

In 1865, Plücker and Hittorf published their classic paper on the appearance of gaseous spectra [33]. They reported that the spectrum of hydrogen could assume a continuous emission as pressures increased: "*Hydrogen shows in the most striking way the expansion of its spectral lines, and their gradual transformation into a continuous spectrum... On employing the Leyden jar, and giving to the gas in our new tubes a tension of about 60 millims, the spectrum is already transformed to a continuous one, with a red line at one of its extremities. At a tension of 360 millims. the continuous spectrum is high increased in intensity, while the red line Hα, expanded into a band, scarcely rises from it*" [33, p. 21–22]. Wüllner quickly confirmed pressure broadening in gaseous spectra [34,35]. Relative to hydrogen, he wrote: "*As the pressure increases, the spectrum of hydrogen appears more and more like the absolutely continuous one of an incandescent solid body*" [35].

During this same period, Frankland [36] and Lockyer made the critical transition of applying line broadening explicitly to the Sun [37]. Much of this discussion was reproduced in Lockyer's text [5, p. 525–560]. They proposed that pressure alone resulted in spectral broadening, excluding any appreciable effects of temperature. This was something which, according to them, had escaped Plücker and Hittorf [33]. They refuted Kirchhoff's solid or liquid photosphere: "*We believe that the determination of the above-mentioned facts leads us necessarily to several important modifications of the received history of the physical constitution of our central luminary — the theory we owe to Kirchhoff, who based it upon his examination of the solar spectrum. According to this hypothesis, the photosphere itself is either solid or liquid, and it is surrounded by an atmosphere composed of gases and the vapours of the substances incandescent in the photosphere... With regard to the photosphere itself, so far from being either a solid surface or a liquid ocean, that it is cloudy and gaseous or both follows both from our observations and experiments*" [37].

Unfortunately, the concept that the spectrum of a gas can be pressure broadened had little relevance to the problem at hand. The line shape was not correct, though this difficulty escaped scientists of this period. The full solar spectrum was not available, until provided by Langley in early 1880's [71–73]. The spectrum of the Sun was not simply broadened, but had the characteristic blackbody appearance, a lineshape that gases failed to reproduce, despite the insistence of Kirchhoff's law to the contrary. In 1897, W. J. Humphreys published his extensive analysis of the emission spectra of the elements [77]. The work only served to re-emphasize that not a single gas ever produced a blackbody spectrum [67–69] through pressure broadening. As a result, the fifth pillar had never carried any real relevance to solar problems.

Hence, astrophysics has had to contend with the inability to generate a Planckian spectrum [67–69] from gases. The spectrum so easily obtained with graphite or soot [61, 65] remained elusive to gaseous solar models, unless recourse was made to a nearly infinite mixture of elemental species and electronic processes [74–76]. As a mechanism, pressure broadening would fall far short of what was required. *A priori*, it shared nothing with the fundamental mechanism existing in graphite and soot, the two best examples of true blackbodies in nature. Consequently, the intriguing discovery of pressure broadening in the 1860's has failed solar science. In reality, the search for the origin of the solar spectrum using gaseous emission spectra has continued to evade astrophysics until the present day, as evidenced by the very existence of The Opacity Project [74, 75].

2.6 The stellar equations of state

Many scientists have not recognized that a slow transformation is taking place in the physical sciences. In large part, this is due to the elegance of the stellar equations of state [15–21] as they continued to evolve from the seminal thoughts of Lane [78], Schuster [79, 80], Very [81], and Schwarzchild [82]. As such, astronomy continues to advocate a gaseous Sun. In doing so, it sidesteps the consequences of solar phenomena and attempts to endow its gaseous models with qualities known only to condensed matter. Simplicity beckons the liquid photosphere through every physical manifestation of its state [57–60]. But, solar physics remains bound by the gaseous plasma.

3 Historical account of the constitution of the Sun

3.1 William Herschel, speculation, and the nature of scientific advancements

Throughout scientific history, the nature of the Sun has been open to changing thought (see Table 1) and, in hindsight, often wild speculation. Even the strangest ideas of our forefathers possess redeeming qualities. It is almost impossible, for instance, to escape the intellectual delight which day-

author	year	sunspots	photosphere	solar body
Thales [5, p. 2]	600 B.C.	?	?	solid
Galileo [101, p. 124]	1612	clouds	fluid	?
Descarte [100, p. 147]	1644	opaque solid mass	fluid	fluid
de la Hire [98, p. 391]	1700	opaque solid mass	fluid	fluid
J. Lalande [98]	1774	opaque solid mass	fluid	fluid
A. Wilson [84]	1774	cavities in photosphere	fluid	dark and solid
W. Herschel [83]	1795	cavities in photosphere	luminous cloud layer	inhabited solid
W. Herschel [88]	1801	cavities in photosphere	luminous cloud/reflective cloud	inhabited solid
F. Arago [89, p. 29]	1848	openings in photosphere	gaseous	solid
J. Herschel [93, p. 229]	1849	cavities in photosphere	luminous cloud/reflective cloud	dark solid
H. Spencer [104, 105]	1858	cyclones	incandescent liquid	gaseous
G. Kirchhoff [43]	1862	clouds	incandescent liquid	solid or liquid
W. Thomson [47]	1862	?	incandescent liquid	incandescent liquid
A. Secchi [95, 96]	1864	openings in photosphere	gaseous with condensed matter	gaseous
J. Herschel [97]	1864	cavities in photosphere	gas?/vapour?/liquid?	dark solid
H. Faye [111, 112, 120]	1865	openings in photosphere	gaseous with condensed matter	gaseous
de la Rue, Stewart, Loewy [133]	1865	openings in photosphere	gaseous with condensed matter	gaseous
Frankland and Lockyer [37]	1865	openings in photosphere	gaseous with condensed matter	gaseous
H. Faye [119]	1872	cyclones	gaseous with condensed matter	gaseous
Modern theory	present	gaseous (magnetic fields)	gaseous	gaseous

Table 1: A partial summary of humanity's concept of the Sun.

dreams of William Herschel's 'solarians' invoke [83]. An inhabited solid solar surface might seem absurd by our standards, but such beliefs dominated a good portion of 19th century thought, at least until the days of Kirchhoff and the birth of solar spectral analysis [30–32, 43]. If Herschel's solarians are important, it is not so much because their existence holds any scientific merit. The solarians simply constitute a manifestation of how the minds of men deal with new information.

As for the concept that the Sun was a solid, the idea had been linked to Thales [5, p. 2], the Greek philosopher, who is said to have pondered upon the nature of the Sun in the 6th century B.C., although no historical evidence of this fact remains [2, p. 81–84]. Lockyer provided a brief discussion of ancient thought on the Sun [5, p. 1–12], in which we were reminded of the words of Socrates that "*speculators on the universe and on the laws of the heavenly bodies were no better than madmen*" [5, p. 5]. Relative to a solid Sun, Herschel did not deviate much from the thoughts of the ancient philosophers whose conjectures were, at times, fanciful [2].

With regard to the photosphere and the "outer layers of the Sun", Herschel placed his distinct mark on solar science. In doing so, he built on the foundation advanced by his predecessor, Alexander Wilson, in 1774 [84]. Herschel wrote: "*It has been supposed that a fiery liquid surrounded the sun, and that, by its ebbing and flowing, the highest parts of it were occasionally uncovered, and appeared under the shape of dark spots; and in that manner successively assumed different phases*" [83, p. 48] ... "*In the instance of our large spot on the sun, I concluded from the appearances that I viewed the real solid body of the Sun itself, of which we rarely see more than its shining atmosphere... The luminous shelving sides of a spot may be explained by a gentle and gradual removal of the shining fluid, which permits us to see the globe of the Sun*" [83, p. 51] ... "*The Sun, viewed in this light, appears to be nothing else than a very eminent, large, and lucid planet, evidently the first, or in strictness of speaking, the only primary one of our system; others being truly secondary to it. Its similarity to the other globes of the solar system with regard to its solidity, its atmosphere, and its diversified surface; the rotation upon its axis, and the fall of heavy bodies, lead us to suppose that it is most probably also inhabited, like the rest of the planets, by being whose organs are adapted to the peculiar circumstances of that vast globe*" [83, p. 63].

Herschel believed that the Sun was a solid globe surrounded by a photosphere made from an elastic fluid which was responsible for light production: "*An analogy that may*

be drawn from the generation of clouds in our own atmosphere, seems to be a proper one, and full of instruction. Our clouds are probably decompositions of some of the elastic fluids of the atmosphere itself, when such natural causes, as in this grand chemical laboratory are generally at work, act upon them; we may therefore admit that in the very extensive atmosphere of the sun, from causes of the same nature, similar phaenomena will take place; but with this difference, that the continual and very extensive decomposition of the elastic fluids of the sun, are of a phosphoric nature, and attended with lucid appearances, by giving out light" [83, p. 59].

Though Herschel first described an inhabited star in 1795, he soon discovered infrared radiation [85–87] and realized that the Sun would provide an uncomfortable setting for its population. In a valiant attempt to save his solarians in 1801, Herschel advanced that the luminous layer of the photosphere, floating like a cloud above the solid solar surface, was positioned beyond an inferior reflective cloud which could channel the heat of the photosphere away from the inhabitants of the Sun [88]. Herschel incorporated a new fact, the discovery of infrared radiation [85–87], with a new concept, the reflective layer [88], in order to salvage an existing theory, the inhabited solid Sun [83]. A study of Herschel reminds us that theories are able to undergo many alterations in order to preserve a central idea, even if the sum of new facts has, long ago, shattered its foundation.

3.2 Alexander Wilson's queries and conjectures

It is noteworthy that, unlike William Herschel, Alexander Wilson, in 1774 (see Table I), displayed uncharacteristic caution for speculation. In elucidating his ideas about the constitution of the Sun, the great astronomer placed the entire text in a section devoted to "*Queries and Conjectures*" [84, p. 20–30]. In fact, he dismissed much of the work of his predecessors as hypotheses without sound scientific basis. He was cautious to highlight the speculative nature of his theory on the constitution of the Sun when he wrote: "*When we consider, that the solar spots, some of whose properties have just now be enumerated, are so many vast excavations in the luminous substance of the Sun, and that, wherever such excavations are found, we always discern dark and obscure parts situated below; is it not reasonable to think, that the great and stupendous body of the Sun is made up of two kinds of matter, very different in their qualities; that by far the greater part is solid and dark; and that this immense and dark globe is encompassed with a thin covering of that resplendent substance, from which the Sun would seem to derive the whole of its vivifying heat and energy? And will not this hypothesis help to account for many phaenomena of the spots in a satisfactory manner? For if a portion of this luminous covering were by means displaced, so as to expose to our view a part of the internal dark globe, would not this give the appearance of a spot?*" [84, p. 20]. He continues: "*And from this may we not infer, that the luminous matter gravitates, and is in some degree fluid...*" [84, p. 22]. Wilson brought forth a solid solar body surrounded by a gaseous or liquid photosphere. He was well aware of the limitations of his own knowledge relative to the photosphere, stating that: "*we may never have a competent notion of the nature and qualities of this shining and resplendent substance...*" [84, p. 21]. Wilson was prudent in the manner by which he proposed new ideas. He closed his address by stating with respect to "*many such other questions, I freely confess, that they far surpass my knowledge*" [84, p. 30]. At the same time, Wilson wrote his "*Queries and Conjectures*" precisely because he realized that they formed a basis for further discovery and questioning. In a field as complex as astronomy, devoid of direct contact with the subject of its attention, mankind could adopt no other logical course of action.

3.3 François Arago, John Herschel, and the constitution of the Sun in the mid-1800's

By the middle of the 19th century, there seemed to have evolved both a popular conception of the Sun and a more "scientific" outlook. François Arago [89, 90], the premier astronomer in France during this period, shed light on the growing divide between popular thought and professional astronomy. He discussed the constitution of the Sun in these terms: "*Many conjectures have been offered in explanation of these spots. Some have supposed that the Sun, from which so vast a quantity of light and heat is incessantly emanating, is a body in a state of combustion, and that the dark spots are nothing else than scoriae floating on its surface. The faculae, on the contrary, they suppose due to volcanic eruptions from the liquified mass. The grand objection to this hypothesis is, that it does not suffice to explain the phenomenoa: it has not obtained admission among astronomers. The opinion most in favor in the present day, regards the Sun consisting of an obscure and solid nucleus, enveloped by two atmospheres — the one obscure, the other luminous. In this case, the appearance of the spot is explained by ruptures occurring in the atmosphere, and exposing the globe of the Sun to view...*" [89, p. 29].

Arago's position constituted essentially a restatement of William Herschel [88]. Only the solarians seemed to have disappeared and the inner atmosphere became obscure, rather than reflective. In order to strengthen his position, Arago then added: "*This opinion, however strange it may appear, has the advantage of perfectly explaining all the phenomena, and it acquires a high degree of probability from the consideration, that the incandescent substance of the Sun cannot be either a solid or a liquid, but necessarily a gas*" [89, p. 29]. Arago justified his position for a gaseous photosphere, well ignorant of the discoveries to come, both of his own time and in the years to follow. He stated: "*It is an established fact that rays of light, issuing from a solid or liquid sphere in a state of incandescence, possess the properties of polarization, while those emanating from incandescent gases are devoid of them*" [89,

p. 29]. He immediately emphasized that polarization experiments support this position affording "*proof that the light of the Sun's edges is as intense as that at its center*" [89, p. 29]. Further, "*But from the fact that the light from the edges of the Sun's disk is as intense as that from the center, there follows another consequence; namely, that the Sun has no other atmosphere outside the luminous one; for otherwise the light of the edges, having a deeper layer to penetrate, would be found more weakened*" [89, p. 29].

Of course, François Arago was incorrect in stating that "*light of the Sun's edges is as intense as that at its center*" [85, p. 29]. In fact, the converse was first observed in the days of Galileo [7, p. 274]. Arago's contemporary, Sir John Herschel, wrote: "*The deficiency of light at the borders of the visible disc is in fact so striking, whether viewed through coloured glasses or without their intervention, by projecting its image through a good achromatic telescope on white paper, that it seems surprising it should ever have been controverted*" [91, p. 434]. Yet, Arago had the notion that a difference in path length through gas would account for differences in observed solar brightness. This was not far removed from the modern concept of optical depth which explained the same phenomenon [79–82, 92]. However, in this instance, it is the light visualized from the center of the Sun which is from deeper, and therefore warmer, regions. For modern solar astronomy, differing path lengths into the Sun permit the sampling of warmer areas. In any case, Arago's arguments, relative to polarization as restated in his *Popular Astronomy* [90, p. 457], would be eventually refuted (see below).

As for John Herschel [91, 93, 94], over most of the course of his life, he viewed the constitution of the Sun through the eyes of his father, William: "*But what are the spots? Many fanciful notions have been broached on this subject, but only one seems to have any degree of physical probability, viz. that they are the dark, or at least comparatively dark, solid body of the Sun itself, laid bare to our view by those immense fluctuations in the luminous regions of its atmosphere, to which it appears to be subject*" [93, p. 229]. He stated that the "*more probable view has been taken by Sir William Herschel, who considers the luminous strata of the atmosphere to be sustained far above the level of the solid body by a transparent elastic medium, carrying on its upper surface... a cloudy stratum which, being strongly illuminated from above, reflects a considerable portion of the light to our eyes, and forms a penumbra, while the solid body shaded by the clouds, reflects none*" [93, p. 229]. The same citation can be found in the 10th edition of his work, published in 1869 [94, p. 314–315]. However, in 1864, along with Father Angelo Secchi [95, 96], John Herschel became one of the first professional astronomers to advance the concept that the Sun was gaseous when discussing sunspots in April of that year: "*while it agrees with that of an aggregation of the luminous matter in masses of some considerable size, and some degree of consistency, suspended or floating at a level determined by their ... gravity in a non-luminous fluid; be it gas, vapour, liquid, or that intermediate state of gradual transition from liquid to vapour which the experiments of Gagniard de la Tour have placed visibly before us*" [97]. In so doing, John Herschel was the first to propose that critical phenomena [26–29] may be important in understanding the structure of the Sun [57]. Oddly, he did not deem these ideas of sufficient merit to modify his popular text. In a public sense, John Herschel remained faithful to his father, even though nearly seventy years had elapsed in the "progress" of science.

3.4 Early thoughts of a fluid Sun

Unlike Alexander Wilson [84] and William Herschel [83, 88], who both advocated a solid solar body, the French astronomer Joseph Jérôme Le Français de Lalande thought that the Sun was a fluid. In his *Abrégé d'astronomie* of 1774 [98], Lalande reiterated the sentiment of his French predecessor, M. de la Hire. In 1700 and 1702, de la Hire stated that a sunspot was most likely the result of "*protrusion of a solid mass, opaque, irregular, swimming in the fluid material of the Sun, in which it sometimes dove entirely*" [98, p. 391]. René Descartes [99, 100] expressed essentially the same ideas in his *Principia Philosophiae*, published in 1644 [100, p. 147–152]. Descartes' contributions were outlined in Karl Hufbauer's classic text [14, p. 21].

Lalande also described how Galileo and Johannes Hevelius viewed the Sun as a fluid: "*Galileo, who was in no manner attached to the system of incorruptibility of the heavens, thought that Sun spots were a type of smoke, clouds, or sea foam that forms on the surface of the Sun, and which swim on an ocean of subtle and fluid material*" [98, p. 390–391]. In 1612, Galileo wrote: "*... I am led to this belief primarily by the certainty I have that that ambient is a very tenuous, fluid, and yielding substance from seeing how easily the spots contained in it change shape and come together and divide, which would not happen in a solid or firm material*" [101, p. 124]. Galileo differed from Lalande in advancing that sunspots were gaseous or cloudy versus solid [101, p. 98–101]. But, Galileo was not attached to this aspect of his work: "*for I am very sure that the substance of the spots could be a thousand things unknown and unimaginable to us, and that the accidents that we observed in them -their shape, opacity, and motion- being very common, can provide us with either no knowledge at all, or little but of the most general sort. Therefore, I do not believe that the philosopher who was to acknowledge that he does not and cannot know the composition of sunspots would deserved any blame whatsoever*" [101, p. 98]. It was the act of locating the spots on, or very close to, the surface of the Sun, that Galileo held as paramount [101, p. 108–124]. Thus, Galileo refuted Scheiner: "*I say that for the present it is enough for me to have demonstrated that the spots are neither stars, nor solid matters, nor located far from the Sun, but that they appear and disappear around it in a manner not dissimilar to*

that of clouds" [101, p. 294–295]. Scheiner, Galileo's constant detractor, believed that special stars strangely coalesced to create sunspots [101, p. 98].

3.5 Kirchhoff, Magnus, Kelvin, and the liquid photosphere

In 1862, Gustav Kirchhoff elucidated the idea of a solid or liquid photosphere: "*In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature. This supposition is in accordance with Laplace's celebrated nebular-theory respecting the formation of our planetary system*" [43, p. 23]. Kirchhoff explained how the Sun, like the planets, was formed through contraction. The Sun remained at the temperature of "white heat" as a result of its greater mass. Kirchhoff cited Arago extensively and was well aware of the work on sunspots by Alexander Wilson. Since the photosphere acted on the body of the Sun, Kirchhoff argued that it must also be heated to the point of incandescence. Relative to the constitution of the Sun, Kirchhoff's entire driving force was the solar spectrum itself. The argument must be echoed, even in the present day.

Unfortunately, it was in speaking of sunspots that Kirchhoff confused the issue: "*But the phenomena exhibited by the solar spots, for whose benefit the hypothesis of a dark solar nucleus was started, may, I believe, be explained more completely and more naturally by help of the supposition concerning the constitution of the sun, which the consideration of the solar spectrum has led me to adopt*" [43, p. 26]. Kirchhoff then advanced that sunspots were the results of layers of clouds which cut off the heat emitted by the incandescent surface of the Sun. Kirchhoff's thoughts were reminiscent of Galileo's [101, p. 98–101], a point not missed by Secchi [3, p. 16], and Faye [5, p. 51–61]. Therefore, Alexander Wilson's cavities were replaced by clouds. Kirchhoff invoked Secchi's work and convection currents to explain why sunspots appear only at certain latitudes and tried to bring understanding to the origin of faculae. This entire portion of the text was somewhat nebulous in logic for a man like Kirchhoff. It would undermine his idea that the photosphere must be solid or liquid based on its continuous spectrum [43].

As an expert in thermal emission, Kirchhoff rapidly objected to Arago's polarization arguments against the liquid. Emphatically, he maintained that Arago's "*statement that incandescent gas is the only source of non-polarized light, is, however, incorrect, for Arago himself mentions that the common luminous gas-flame emits perfectly unpolarized light; and the light in this case is almost entirely caused not by glowing gas, but by incandescent particles of solid carbon which are liberated in the flame. An incandescent haze consisting of solid or liquid particles must act in a manner precisely similar to such a flame*" [43, p. 30]. Kirchhoff further explained that a liquid Sun, whose seas are in continuous motions, would emit light from its surfaces in different directions with respect to our eyes. This destroyed any polarization. The argument was a powerful one, but as will be seen below, it was Kirchhoff's explanation of sunspots which his contemporaries, Secchi and Faye, would reject. In so doing, they would dismiss Kirchhoff's entire vision for the constitution of the Sun. This move on their part reflected, perhaps, their all too hasty conclusions with regards to thermal emission. The error continued to this day.

Heinrich Gustav Magnus [102] also believed that the Sun was a liquid. He was a great supporter of Kirchhoff [43]. On July 11th, 1861, he delivered Kirchhoff's memoire on the chemical constitution of the Sun's atmosphere before the Berlin Academy [103, p. 208]. Magnus demonstrated that the addition of caustic soda (sodium hydroxide) to a non-illuminating gaseous flame generated a tremendous increase in its luminosity [102]. He noted the same effect for the salts of lithium and strontium. In 1864, according to Magnus: "*These studies demonstrate that gaseous bodies emit much less heat radiation than solid or liquid bodies; and that, by consequence, one cannot suppose that the source of solar heat resides in a photosphere composed of gas or vapours*" [102, p. 174]. Magnus' argument was powerful and, for the next 50 years, it continued to impact the constitution of the Sun. It was because of Magnus that photospheric theory would preserve some aspects of condensed matter well into the beginning of the 20th century. It would eventually take the theoretical arguments of men like Schuster [79,80], Very [81], Schwarzschild [82], Eddington [51], and Milne [92] to finally set aside Magnus' contributions [102] and cast the concept of condensed matter out of the photosphere [43].

Kirchhoff liquid Sun was also echoed by William Thomson himself. Lord Kelvin states: "*It is, however, also possible that the Sun is now an incandescent liquid mass, radiating away heat, either primitively created in his substance, or, what seems far more probable, generated by the falling in of meteors in past times, with no sensible compensation by a continuance of meteoric action*" [47]. By the time these words were written, Thomson no longer believed that the Sun could replenish its energy with meteors and wrote: "*All things considered, there seems little probability in the hypothesis that solar radiation is at present compensated, to any appreciable degree, by heat generated by meteors fallings in; and, as it can be shown that no chemical theory is tenable, it must be concluded as most probable that the Sun is at present merely an incandescent liquid mass cooling*" [47]. In the same paper, Thomson discussed Helmholtz' contraction theory, as an extension, it seemed, of the meteoric hypothesis [47]. The contraction and meteoric models of energy generation would eventually prove to be unsound. But, for the

time being, Thomson continued to view the Sun as liquid in nature, as did Kirchhoff and Magnus.

At the same time, it is ironic how Kirchhoff, through his law of thermal emission, unknowingly provided for astrophysics the very basis for the downfall of his liquid model. Currently, the entire concept of a gaseous Sun rests on the presumed validity of Kirchhoff's formulation. Nonetheless, early gaseous models of the Sun always placed either solid or liquid constituents in the region of the photosphere, as shall soon be outlined. Not until the early 20th century would the Sun become fully divested of condensed matter. In so doing, astrophysics would endow the gaseous plasma with emission properties it failed to possess on Earth. Regrettably, few of Kirchhoff's contemporaries supported his idea that the Sun was a liquid. Visual observations, and the view that Kirchhoff was an outsider to astronomy, would become ruinous to his model. Critical temperatures [28] also dictated that the Sun was simply too hot to allow this phase. Spectroscopic evidence became of secondary importance and the journey to a gaseous Sun formally began.

4 On to a gaseous Sun

4.1 Men, ideas, and priority

Throughout the history of astronomy, there is perhaps no more controversial figure than Herbert Spencer. As an independent philosopher, not formally trained in science, he became the first to advance that the interior of the Sun was completely gaseous [104–106]. He was also a staunch supporter of evolution and elucidated the concept of "*survival of the fittest*" [107]. In academic circles, Spencer was widely criticized for the views he held, both in ethics and in sociology [108]. By his supporters, he seemed highly admired [108] and compared to other polymaths including the likes of Goeth, Humbolt, and Whewell [103, p. 198]. Unfortunately, many of Spencer's social thoughts were unfounded and promoted concepts of imperialistic superiority and outright discrimination [107, p. 481–483]. His contributions on the constitution of the Sun [104, 105] were essentially ignored by professional astronomy, even though he corresponded with Sir John Herschel and Sir George Airy, the Astronomer Royal [106]. In addition, Spencer was a close friend of the great physicist John Tyndall who became, in like manner, a prominent evolutionist [106]. Spencer's political and social views were so counter to those espoused by men of the period that he remained ever outside the mainstream of astronomy.

Spencer eventually argued for priority over Hervé Faye with respect to his ideas of a gaseous Sun [105]. His defense was in response to review articles by Norman Lockyer published in the magazine *The Reader* [109, 110], about the Frenchman's *Comptes Rendus* papers [111, 112]. Nine years later, Lockyer reprinted these articles in his classic text [5, p. 44–62], without reference to Spencer's letter [105]. In doing so, Lockyer approached misconduct. He added a footnote crediting Balfour Stewart and Gustav Kirchhoff for a thermodynamic argument which the record well demonstrated was first expounded in Spencer's letter, as will be discussed in Section 4.6 [105]. But since Lockyer was the cause of Spencer's 1865 letter [105], he could not have been unaware of its contents.

Bartholomew advanced a somewhat disparaging analysis of Spencer's contributions to solar physics [106]. He attempted to justify Spencer's rejection by professional astronomy. Though he gave Spencer qualities, he charged him with being simply an amateur, a surprisingly desultory reader, and of incorporating in his own writings facts and ideas acquired in other ways [106]. He even accused Spencer with making the Nebular hypothesis the starting point of his discussion, justifying the same behavior by men like Kirchhoff and Faye as merely supportive and confirmatory [106, p. 22]. Though Bartholomew brought forth several other reasons why Spencer was ignored, many of which were perhaps valid, his central argument was summarized as follows: "*Rather, at the mid-nineteenth century a criterion of acceptability for scientific pronouncements was beginning to emerge that was linked to the notion of professionalism; only those who had credentials in their subject through training and research could expect to have their speculative theories taken seriously. As this standard gradually asserted itself, Spencer's work in astronomy lost much of its claim for attention*" [106, p. 21]. This aspect of 19th century thought, beginning to permeate science in Spencer's day, had also been proposed while discussing Robert Chambers' *Vestiges on the Natural History of Creation* which was one of the first works on evolutionary reasoning: "*the reaction to Vestiges was not simply a profession of empiricism: it was an attempt to restrict the privilege of theoretical speculation to a small circle of recognized researchers*" [113, p. 22].

Relative to the Sun, a review of the documents of the period showed no more theoretical brilliance in the works of Secchi [95, 96, 114–118] and Faye [109–112, 119, 120] than in those of Spencer [104, 105]. This was reality, despite the fact that Spencer was charged with being ill-trained in thermodynamics, astronomy, and mathematics [106]. While Secchi was a magnificent observational astronomer [3], all three men were profoundly mistaken in many of their ideas regarding the Sun and sunspots. Furthermore, in light of modern analysis, their differences hinged on the trivial. Few of the early works of either Secchi or Faye were mathematical in nature [95, 96, 109–112, 114–120].

The nature of sunspots had immediately become a focus of contention between Spencer [105] and Faye [120]. In fact, Secchi and Faye would criticize Kirchhoff on the same subject, although they were far from being his equal in theoretical prowess. In *Comptes Rendus*, the battle between Faye and Kirchhoff on sunspots was protracted, extensive [121–126], and would yield many of the modern ideas for a gaseous Sun. Faye and Secchi's defense against Kirchhoff was some-

what justified, relative to sunspots not resting as clouds above the photosphere. But they did not sufficiently appreciate the importance of the German's arguments for condensed matter [43]. For many decades, the contributions of these two men, on the constitution of the Sun, were highly cited and praised. Spencer, their British colleague, continued to be essentially ignored [106].

Consequently, had the scientific community merely erected a means of self-promotion and preservation, with respect to theoretical speculation, by rejecting Spencer's work? This is unlikely to be the only explanation. It was obvious that many despised Spencer's social, ethical, and evolutionary thoughts. Competitive pressures must also have been involved. Hervé Faye clearly became acquainted with Spencer's work, given the three articles presented in *The Reader*. Still, the Frenchman long delayed to cite Spencer. Yet, it was unlikely that mere "scientific exclusivity" could account for Faye's and Lockyer's treatment of Spencer, as Bartholomew proposed. Hervé Faye defended religion and argued on moral grounds against the merits of evolution in addressing both science and God in his classic text which emphasized: "*Coeli enarrant gloriam Dei*" [127, p. 1–4]. As such, it appears that Faye consciously refused to confer upon Spencer the credit he deserved. This was especially true given the struggle for priority and Faye's time in history [127, p. 1–4]. The situation was perhaps clearer for Father Secchi. Secchi likewise echoed "*Coeli enarrant gloriam Dei*" [128, p. 1] and, on his deathbed, paraphrased Saint Paul (2 Timothy 4:7–8): "*I have finished my course, I have fought the good fight. Throughout my entire life and in my scientific career, I have had no other goal but the exultation of the Holy Catholic Church, demonstrating with evidence how one can reconcile the results of science with Christian piety*" [128, p. vii]. It must be remembered that, when the Jesuits would be expelled from Rome, Secchi was defended by the world scientific community. Only Secchi, with his assistants, was allowed to remain in the city and continued to work at the Observatory of the Roman College [128, p. xxii-xxiii]. Did Secchi know in advance of Spencer's *Westminster Review* article [104]? In 1869, Secchi had mentioned, with respect to Lockyer, that "*As to what regards his work, I admit that I have knowledge of only those which were published in Comptes Rendus, or in Les Mondes*" [5, p. 500]. The situation is not definitive however, as Secchi does mention his knowledge of the recent work by William R. Dawes in *Monthly Notices* in his first letter [95]. Nonetheless, it was doubtful that the Director of the Observatory of the Roman College knew of Spencer's works when he wrote his key papers of 1864 [95, 96]. The surest evidence was the lack of similarity between the ideas of Secchi [95, 96] and Spencer [104]. Conversely, this was not the case for Faye's classic papers [111,112], including those dealing with the defense of his sunspot theory [119–126]. The problem for Faye would be three fold: 1) extensive scientific similarity, 2) eventual and certain knowledge of Spencer's rebutal letter in *The Reader* [105] and 3) his claim of simultaneous discovery with respect to Secchi, as will be soon discovered. For Faye at least, it is difficult to argue against deliberate scientific disregard relative to Spencer and his ideas.

Relative to issues of faith, it is also notable that many learned men of the period shared Faye's and Secchi's dual affection for religion and science. In fact, even Max Planck would be counted in their company [129]. Bartholomew failed to address any of these points. It is unlikely that the dismissal of Spencer can be solely attributed to his lack of training, amateur status, and "*an attempt to restrict the privilege of theoretical speculation to a small circle of recognized researchers*" [113, p. 22]. The reality remained that some of Spencer's ideas continued to be objectionable (e.g. [107, p. 481–483]) and that the quest for priority was powerful.

Nonetheless, one must question the persistent failure [7, 13, 14] to give Spencer credit for advancing the earliest model of the gaseous Sun. Bartholomew's discussion [106], in trying to justify the past with the privilege of scientific position and "right to speak", did nothing to advance truth. This was especially highlighted, when contrasted with Galileo's free acknowledgement of Benedetto dei Castelli's contributions to the projection of sunspots [101, p. 126]. It was further expounded by the remembrance of Charles' law by Gay-Lussac [49], even though the former had not written a single word and the experiments were done fifteen years earlier. If the name of *Charles' law* exists, it is only because of Gay-Lussac's profound honesty. As such, the refusal to credit Spencer for his contributions should not be justified by modern writers [106], but rather, must be condemned as an unfortunate injustice relative to acknowledging the genesis of scientific ideas [130]. The reality remains that the birth of a gaseous Sun was accompanied by bitter rivalry throughout professional astronomy, much of which was veiled with struggles for priority. In this expanded context, and given his social views, Spencer's isolation was not surprising.

4.2 Herbert Spencer and the nebular hypothesis

In reality, Spencer's contributions were noteworthy for their dramatic departure from the ideas of Herschel and Arago (see Table 1). Much like other works of the period, Spencer's thesis contained significant scientific shortcomings. Still, his writings were on par with those of his contemporaries and were, it appears without question, the first to outline both a gaseous solar body and a liquid photosphere. Spencer advanced this model in an unsigned popular work entitled *Recent Astronomy and the Nebular Hypothesis* published in the *Westminster Review* in 1858 [104]. He began his thesis by imagining a "*rare widely-diffused mass of nebulous matter, having a diameter, say as great as the distance from the Sun to Sirius*" [104, p. 191] and considered that mutual gravitation would eventually result in the "*slow movement of the atoms towards their common center of gravity*" [104, p. 191]. He argued that, as the nebular mass continued to contract, some

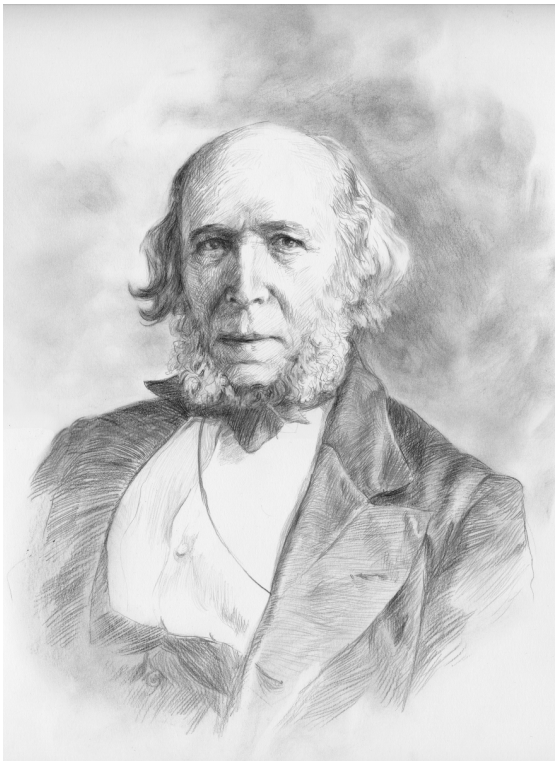


Fig. 1: Herbert Spencer (April 27th, 1820 — December 8th, 1903), was a polymath who advanced the first gaseous model of the Sun, in 1858 [104]. He conceived of a "Bubble Sun", a gaseous interior of variable density surrounded by a fully liquid photosphere. (Drawing by Bernadette Carstensen — used with permission.)

of the internally situated atoms entered into chemical union. With time, as the heat of chemical reaction escaped the nebular mass, the latter began to cool. The binary atoms would then precipitate and aggregate into "*flocculi*" [104, p. 192]. Spencer described how *flocculi* formation resulted in centripetal motion of the nebula and eventually condensed into a larger internal and external aggregate masses. The latter developed into planets and comets. Spencer summarized Laplace's nebular hypothesis as follows: "*Books of popular astronomy have familiarized even unscientific readers with his [Laplace's] conceptions; namely, that the matter now condensed into the solar system once formed a vast rotating spheroid of extreme rarity extending beyond the orbit of Neptune; that as it contracted its rate of rotation necessarily increased; that by augmenting centrifugal force its equatorial zone was from time to time prevented from following any further the concentrating mass, and so remained behind as a revolving ring; that each of the revolving rings thus periodically detached eventually became ruptured at its weakest point, and contracting upon itself, gradually aggregated into a rotating mass; that this like the parent mass, increased in rapidity of rotation as it decreased in size, and where the centrifugal force was sufficient, similarly through off rings, which finally collapsed into rotating spheroids; and that thus out of these primary and secondary rings arose the planets and their satellites, while from the central mass there resulted the Sun*" [104, p. 201].

Spencer succinctly outlined his thoughts on the Sun when he defended himself in *The Reader*. He opened as follows: "*The hypothesis of M. Faye, which you have described in your numbers for January 28 and February 4, is to a considerable extent coincident with one which I ventured to suggest in an article on 'Recent Astronomy and the Nebular Hypothesis,' published in the Westminster Review for July, 1858. In considering the possible causes of the immense differences of specific gravity among the planets, I was led to question the validity of the tacit assumption that each planet consists of solid or liquid matter from centre to surface. It seemed to me that any other internal structure, which was mechanically stable, might be assumed with equal legitimacy. And the hypothesis of a solid or liquid shell, having its cavity filled with gaseous matter at high pressure and temperature, was one which seemed worth considering, since it promised an explanation of the anomalies named, as well as sundry others*" [105]. He continued: "*The most legitimate conclusion is that the Sun is not made up of molten matter all through; but that it must consist of a molten shell with a gaseous nucleus. And this we have seen to be a corollary of the Nebular Hypothesis*" [105].

Throughout the article in *The Reader*, Spencer cited extensively from his prior work [104]. The resemblance to Faye's 1865 papers [111, 112] was difficult to justify as coincidental. Spencer argued strongly for the existence of convection currents within the Sun: "*... hence an establishment of constant currents from the center along the axis of rotation towards each pole, followed by a flowing over of accumulation at each pole in currents along the surface to the equator; such currents being balanced by the continual collapse, towards the center, of gaseous matter lying in the equatorial plane*" [105]. The presence of convection currents was to become a central aspect of Faye's model. Nonetheless, Spencer was arguably one of the first to invoke true convection currents within the Sun.

There were several elegant strokes in Spencer's original paper in the *Westminster Review* [104], including his anticipation of the contraction hypothesis which he re-emphasized in *The Reader*: "*Supposing the Sun to have reached the state of a molten shell, enclosing a gaseous nucleus, it was concluded that this molten shell, ever radiating its heat, but ever acquiring fresh heat by further integration of the sun's mass, will be constantly kept up to that temperature at which its substance evaporates*" [105]. He advanced two strata of atmosphere above the molten solar surface, the first "*made up of sublimed metals and metallic compounds*" and the second of "*comparatively rare medium analogous to air*" [105].

Spencer was concerned with the specific gravity of the sun, insisting "*but the average specific gravity of the Sun is about one*" [105]. He ventured: "*The more legitimate conclu-*

sion is that the sun's body is not made up of molten matter all through, but that it consists of a molten shell with a gaseous nucleus... the specific gravity of the Sun is so low as almost to negative the supposition that its body consists of solid or liquid matter from the center to surface, yet it seems higher than is probable for a gaseous spheroid with a cloudy envelope" [105]. Spencer reached this conclusion because he considered only the specific gravity of the metals and materials on Earth. He never realized that the Sun was mostly made of hydrogen. As such, given his building blocks, Spencer was left with a gaseous interior. The insight was profound. In fact, the objection which Spencer made, with respect to the improbability of a gaseous spheroid, would be repeated by the author, before he became acquainted with Spencer's writings [57].

Specific gravity has become a cornerstone of the modern liquid metallic hydrogen model of the Sun [57–60]. At the same time, science must marvel at the anticipation which Spencer gave of the current gaseous models of the Sun when he wrote: "*...but that the interior density of a gaseous medium might be made great enough to give the entire mass a specific gravity equal to that of water is a strong assumption. Near its surface, the heated gases can scarcely be supposed to have so high a specific gravity, and if not, the interior must be supposed to have a much higher specific gravity*" [105]. This is precisely what is assumed by astronomy today, as it sets the photospheric density to $\sim 10^{-7}$ g/cm^3 and that of the solar core to $\sim$150 g/cm^3 [57]. With respect to convection currents and intrasolar density, it could be argued that Spencer led astrophysical thought.

Spencer closed his defense by restating his theory of sunspots. He initially advanced that the spots were essentially cyclones and credited John Herschel with the idea [105]. He then stated that cyclones contained gases and that the effects of refraction could account for their dark appearance. Spencer would modify his idea over time, but he continued to focus on cyclones. His conjectures regarding sunspots would have no redeeming features for the current understanding of these phenomena. As such, suffice it to re-emphasize the novelty of Spencer's *Bubble Sun* as a significant departure from the solid model of the period, with the introduction of convection currents and arguments regarding internal solar density.

4.3 Angelo Secchi and the partially condensed photosphere

Angelo Secchi [3] first outlined his ideas regarding the physical constitution of the Sun in the *Bullettino Meteorologico dell' Osservatorio del Collegio Romano* in two 1864 manuscripts [95, 96]. John Herschel followed suit in April of the same year [97]. Secchi's January work, represented a gentle rebuttal of Gustav Kirchhoff, initially relative to sunspots: "*Signor Kirchoff rejects both the theory of Herschel and that of Wilson. We will first permit ourselves the observation that it is one thing to refute Herschel's theory, and quite another to refute Wilson's, and that when the first is laid to rest, the second one hardly collapses*" [95]. Secchi also disagreed with Kirchhoff relative to thermal emission, disputing that all objects at the same temperature produce the same light: "*Kirchoff relies greatly on the principle that all substances become luminous at the same temperature in order to prove that the core of the sun must be as bright as the photosphere. Here it seems to us that two quite different matters have been conflated: that is, the point at which bodies begin to excite luminous waves capable of being perceptible to the eye, and the fact that all [substances] at the same temperature should be equally luminous. We can accept the first of these propositions, and wholly reject the second. In furnaces we see gases of entirely different luminosity from that of solids, and the strongest [hottest] flame that is known — that is, that of the oxyhydrogen blowpipe — is it not one of the least luminous?*" [95]. In this respect, Secchi was actually correct, as Kirchhoff had inappropriately extended his law to liquids and gases. Secchi realized that gases could not follow Kirchhoff's supposition. This was a rare instance in the scientific literature where the conclusions of Kirchhoff were brought into question. Secchi also expounded on his theory of the

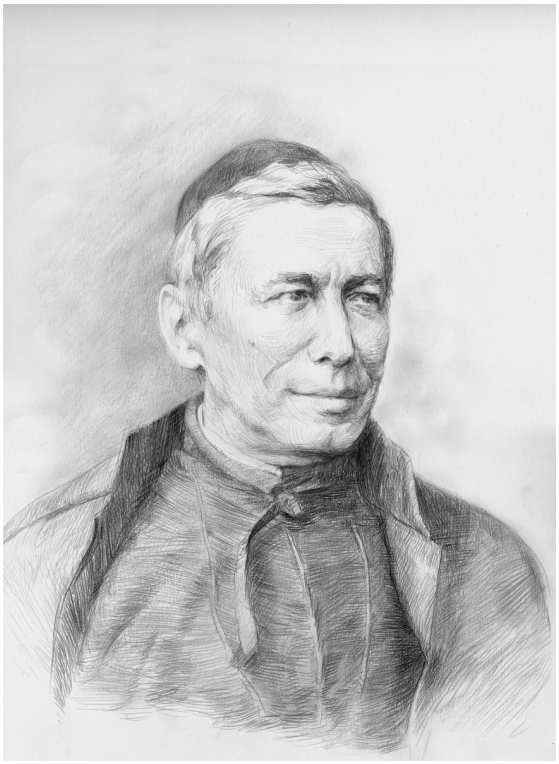


Fig. 2: Father Angelo Secchi, S.J. (June 29th, 1818 — February 26th, 1878), was one of the foremost solar astronomers of his day and the Director of the Observatory of the Roman College. In 1864, Secchi advanced a solar model wherein the photosphere was formed of solid or liquid particulate matter floating on the gaseous body of the Sun [95, 96]. (Drawing by Bernadette Carstensen — used with permission.)

Sun in his classic text [95, p. 37]. Nonetheless, considering Secchi's position, his first article displayed a certain sternness with respect to Kirchhoff, closing with the words: "*We wanted, therefore, to say these things less to object to such a distinguished physicist, than to prevent science from taking a retrograde course, especially since history shows that persons of great authority in one branch of knowledge often drag along, under the weight of their opinion, those who are less experienced, even in matters where their studies are not sufficiently deep and where they should not have such influence*" [95]. Secchi appeared to be arguing, much like Bartholomew [106], that astronomy had become too specialized for the non-professional, even if represented by Kirchhoff himself.

The heart of Secchi's conception of the Sun was outlined in his November 1864 paper [96]. Secchi was concerned with the physical appearance of the solar surface: "*The grid-like solar structure seemed to us to offer nothing regular in those parts of the disc that are continuous, and thus the term granular appears very appropriate. Nevertheless, in the vicinity of the sunspots, that of willow leaf remains justified, because we actually see a multitude of small strips which terminate in rounded tips, and which encircles the edge of the penumbra and of the nucleus, resembling so many elongated leaves arranged all around. The granular structure is more visible near the spots, but it is not recognizable in the faculae; these present themselves like luminous clusters without distinguishable separation, emitting continual light without the interruption of dots or of that black mesh*" [96]. He then clarified his model of the solar photosphere: "*Indeed this appearance suggests to us what is perhaps a bold hypothesis. As in our atmosphere, when it is cooled to a certain point, there exists a fine substance capable of transforming itself in fine powder and of forming clouds in suspension, (water transforming into so-called 'vesicular' vapor or into small solid icicles), so in the enflamed solar atmosphere there might be an abundance of matter capable of being transformed to a similar state at the highest temperatures. These corpuscles, in immense supply, would form an almost continuous layer of real clouds, suspended in the transparent atmosphere which envelopes the sun, and being comparable to solid bodies suspended in a gas, they might have a greater radiant force of calorific and luminous rays than the gas in which they are suspended. We may thus explain why the spots (that are places where these clouds are torn) show less light and less heat, even if the temperature is the same. The excellent results obtained by Magnus, who has proved that a solid immersed in an incandescent gas becomes more radiant in heat and light than the same gas, seem to lend support to this hypothesis, which reconciles the rest of the known solar phenomena*" [96]. Secchi's model differed from Spencer's [104, 105] in that his photosphere was not a continuous layer of liquid. Rather, Sechhi's Sun was essentially gaseous throughout. In his photosphere, solid matter was suspended within the gas. Secchi adopted this model as a result of his visual observations and of Magnus' work on the thermal emission of caustic soda in the transparent gas flame [102]. In this regard, Secchi demonstrated a relatively good understanding of thermal emission.

Over the years, Secchi refined his model of the Sun, but the discussions would be highly centered on the nature of Sun spots. Secchi was a prolific author with more than 800 works to his name [128, p. xvi]. A partial listing of these, compiled at his death, included more than 600 publications [128, p. 95–120]. By necessity, the focus will remain limited to only five of his subsequent contributions on the Sun [114–118].

In the first of these publications [114], Secchi examined sunspots and largely confirmed Wilson's findings [84] that sunspots represented depressions on the solar disk. For both Secchi and Faye, this became a key objection to Kirchhoff's "*cloud model*" of sunspots [43].

In the second article, published in 1868 [115], the astronomer was concerned with the observation of spectral lines in the corona, but he concluded with a defense of the gaseous Sun. Secchi referred to a "*famous objection*" against his model, but never named the source. In actuality, for Secchi, the source of the objection must have been Kirchhoff's *Comptes Rendus* article, which appeared the previous year: "*From the relation which exists between the emissive and absorptive power of bodies, it results in an absolutely certain manner, because in reality the light emitted by the solar nucleus is invisible to our eye, this nucleus, whatever its nature may be, is perfectly transparent, in such a manner that we would visualize, through an opening situated on the half of the photosphere turned in our direction, through the mass of the solar nucleus, the internal face of the other half of the photosphere, and that we would perceive the same luminous sensation as if there was no opening*" [121, p. 400]. Kirchhoff's objection was almost identical to that first leveled by Spencer in 1865 [105, p. 228]: "*But if these interior gases are non-luminous from the absence of precipitated matter must they not for the same reason be transparent? And if transparent, will not the light from the remote side of the photosphere, seen through them, be nearly as bright as that from the side next to us?*" Kirchhoff had strong ties with Guthrie, Roscoe, and the English scientific community. In addition, in light of the previous incident between Kirchhoff and Stewart on priority in thermal emission [61, 138] it is difficult to imagine that the German scientist was unaware of Spencer's work. Two years had already passed.

In response to Kirchhoff, Secchi stated: "*The objection consisted in holding that, if Sun spots were openings in the photosphere, one should be able to see through a gaseous solar mass the luminous photosphere on the other side: as a result, Sun spots would be impossible, since they are not luminous, but black*" [115]. Secchi advanced two lines of defense: "*1) that sunspots, even in their nucleus, are not deprived of light and 2) that for the entire solar mass to be able to produce an absorption capable of preventing the visualiza-*

tion of the other side, it suffices that the interior of the Sun possess an absorbing power identical to its external atmosphere" [115]. Here was perhaps the conclusion of one of the first discussions concerning internal stellar opacity. It reflected why Spencer's complaint was central to the history of astronomy.

Secchi's third work in this series [116] was surprising for two reasons. First, Secchi described that he "*even believes he has seen traces of water vapour in the Sun, especially near the sunspots*" [116, p. 238]. Secondly, and most importantly, Secchi appealed to the French scientific community and to Mr. Sainte-Claire Deville to work on observing the incandescent light emitted by hydrogen under conditions of high pressure [116]. Sainte-Claire Deville immediately followed Secchi's letter with an affirmative response. Secchi thus highlighted the importance of line broadening in hydrogen [33–37] for astrophysical thought [116, p. 238].

In the fourth work of this series, Secchi once again argued that "*sunspots are cavities in the photosphere in whose interior the absorbing layer is thicker*" and continues that "*the brilliant lines that often traverse their nucleus could well be the direct lines of that gas which I have signaled constitutes the gaseous mass of the interior of the Sun*" [117, p. 765]. Secchi was completely mistaken, as these lines do not originate from inside the solar body. His 1869 argument [117] was also counter to that which he already outlined when speaking on stellar opacity a year earlier [115].

In the final work of interest, Secchi described four possible aspects of the chromosphere including: "*The first aspect is one of a layer clearly terminated, as would be the free surface of a liquid... sometimes, especially in the region of faculae, the surface is diffuse*" [118, p. 827]. Secchi completed his 1872 work with a detailed visual description of prominences.

Secchi also entered into a prolonged confrontation in *Comptes Rendus*, initiated by Lockyer, over the constitution of the Sun (reprinted in [5, p. 500–515]). The arguments were spectroscopic in nature and focused on the photosphere, the reversing layer, and the chromosphere. The rivalry, surrounding the gaseous models, had become intense.

In summary, a detailed review of Secchi's work reveals that he was truly an "observational astronomer". Though his initial contributions on the Sun were devoid of mathematical arguments, he displayed a keen sense of deduction, a broad scientific knowledge, and a profound honesty. Unlike Spencer [104, 105], Secchi did not bring to prominence the presence of convection currents inside his gaseous Sun. He based his solar model on the appearance of the solar surface and the work of Magnus [102]. Secchi opposed Kirchhoff [43] on the appearance of sunspots, correctly arguing for Wilson's cavities [84]. Secchi also disputed Kirchhoff's law [30–32] as experimentally unfounded relative to gases [95]. In his book, Secchi provided a discussion of thermal radiation [3, p. 311–319], reminding us of the work of Melloni who demonstrated that: "*different substances possess a particular and elective absorbing force, each of which acts on different heat rays, absorbing some while permitting others to pass, much like colored media acts on white light*" [3, p. 311]. Herein lays Secchi's objection to the universality of Kirchhoff's formulation [30–32]. He recognized the emphasis of his day on line broadening [33–37] and was one of the first to invoke significant stellar opacity [115]. Unfortunately, he advanced seeing water on the solar surface [116, p. 238]. Eventually, mankind would indeed discover water on the Sun [131], but Secchi and his model, by then, would be long forgotten.

4.4 de la Rue, Stewart, Loewy, Frankland, and Lockyer

Shortly after Secchi published his commentaries in *Bullettino Meteorologico* and in *Les Mondes* [95, 96], Warren de la Rue, Balfour Steward, and Benjamin Loewy made their famous report on their theory of sunspots on January 26, 1865. Armed with the sunspot observations of Carrington [132], they expanded on his discoveries [133–137]. Carrington led a tragic life [138, p. 117–128] and was an amateur [13, p. 32]. His observational work, unlike Spencer's ideas, became a cornerstone of astronomy. Presumably, this was because Carrington established the differential rotation of the Sun [132]. He also stayed clear of controversial philosophy and of theorizing on the internal constitution of the Sun. As for de la Rue, Stewart, and Loewy, their contributions with the photoheliograph at Kew were significant. As professional scientists, they ventured into a discussion on the constitution of the photosphere. Historically, their classic paper [133], like Faye's [111, 112], also appeared immediately after the *Les Mondes* translation of Secchi's seminal work [96].

Nonetheless, de la Rue, Stewart, and Loewy were the first [133] to propose that the continuous solar spectrum was consistent with a *fully* gaseous atmosphere. They were quickly endorsed by Frankland and Lockyer who, after believing they had disarmed Kirchhoff, wrote: "*That the gaseous condition of the photosphere is quite consistent with its continuous spectrum. The possibility of this condition has also been suggested by Messrs. De la Rue, Stewart, and Loewy*" [37]. The argument was based on the existence of pressure broadening, observed with hydrogen under conditions of high pressure [37]. It was here that pressure broadening became permanently linked to the gaseous models of the Sun. However, the idea of a fully gaseous photosphere would not truly take hold until much later. For most scientists, the photosphere continued to have at least traces of condensed matter. As for the concept that hydrogen, under pressure, could create a Planckian blackbody spectrum, it was always erroneous. Gases could never produce the required emission [77]. Frankland and Lockyer could not have established this fact with the experimental methods of 1865. They merely observed that the hydrogen lines became considerably broadened, completely unaware of their incorrect lineshape. Irrespective of this shortcoming, the paper by Frankland and

Loewy impacted scientific thought for the rest of the century and became highly cited by the astronomical community. As such, Frankland and Lockyer, along with de la Rue, Stewart, and Loewy who had so magnificently photographed the Sun, hold a preeminent role in the history of solar science [37, 133–137].

Addressing faculae, de la Rue and his team reported: "*It would thus appear as if the luminous matter being thrown up into a region of greater absolute velocity of rotation fell behind to the left; and we have thus reason to suppose that the faculous matter which accompanies a spot is abstracted from that very portion of the sun's surface which contains the spot, and which has in this manner been robbed of its luminosity*" [134]. Based on such observations, they ventured: "*From all of this it was inferred that the luminous photosphere is not to be viewed as composed of heavy solid, or liquid matter, but is rather of the nature either of a gas or cloud, and also that a spot is a phenomenon existing below the level of the sun's photosphere*" [134]. The proposal resembled Secchi's [95, 96]. With these words, Kirchhoff's thermodynamic reasoning, regarding the continuous solar spectrum, became supplanted by visual observations and the Sun adopted the gaseous state.

Given Stewart's earlier conflict with Kirchhoff [61, 139], it would not be unexpected if the Scottish astronomer, at the side of de la Rue and Loewy, had agreed to dispense with Kirchhoff's condensed photosphere [133–135]. However, this was not to be the case. Stewart, a man of strong moral character [140, 141], immediately abandoned de la Rue's gaseous sun, as we will come to discover in Section 4.7.

Beyond Stewart, a historical review of the period reveals that virtually every prominent astronomer voiced public disapproval of Kirchhoff's liquid photosphere. In a real sense, Kirchhoff stood essentially undefended against much of the scientific community. Yet, were the arguments of men like Secchi, Faye, de la Rue, and Lockyer truly sufficient to eventually advance a fully gaseous photosphere? Note in this regard, the *faux pas* by de la Rue, Stewart, and Loewy as to the cause of sunspots in their very next paper: "*the behavior of spots appears to be determined by the behavior of Venus*" [134]. Though Kirchhoff might have misjudged the nature of sunspots, the fault was minor and irrelevant today when compared to the error of assigning an improper phase to the entire Sun. In this respect, Galileo's words in his first letter to Welser come to mind: "*For the enemies of novelty, who are infinite in number, would attribute every error, even if venial, as a capital crime to me, now that it has become customary to prefer to err with the entire world than to be the only one to argue correctly*" [101, p. 89].

4.5 Hervé Faye and loss of the solar surface

Hervé Faye opened his classic presentation on the constitution of the Sun on January 16th, 1865, by stating that the solar phenomena had been well popularized [103]. Therefore, he reduced his historical discussions to the strict minimum and limited himself to the simple analysis of current facts and conjectures [111]. He set the stage by recalling the gaseous envelope and the polarization arguments of Arago [111, p. 92–93]. At the same time, he recognized the importance of Kirchhoff's spectroscopic studies and wrote: "*But incandescent solids and liquids alone give a continuous spectrum, while the gases or the vapors supply but a spectrum reduced to only a few luminous rays*" [111, p. 93]. Faye then argued against Kirchhoff's view of sunspots, as rejected, even by Galileo [111, p. 94]. He proposed that sunspots were produced by clearings in the photosphere, thereby exposing the nucleus of the Sun. Interestingly, Faye argued for the oblateness of the Sun based on the fluidity of the photosphere. Unfortunately for him, the slight oblateness of the Sun [142] supported a condensed photosphere, not one with a gaseous composition [57]. In his seminal communication [111], Faye did not actually advance a complete solution for the nature of the photosphere. He reserved this critical step for his second paper [112].

Fig. 3: Hervé Faye (October 1st, 1814 — July 4th, 1902) was a prominent French astronomer with a distinguished career in science and public service as a minister of education. In early 1865, Faye echoed Secchi's solar model wherein the photosphere was formed of solid or liquid particulate matter floating on the gaseous body of the Sun [111, 112]. (Drawing by Bernadette Carstensen — used with permission.)

Throughout his first work [111], Faye cited many notable figures, but failed to mention either Magnus or Spencer and,

more importantly, Secchi's model [111]. Faye studied under the tutelage of François Arago who, as discussed in Section 3.3, visualized a divide between professional astronomy and popular thought, even in the first half of the 19th century. As such, Bartolomew's arguments for the failure to cite Spencer might be given some weight [106]. But what of Faye's failure to mention Secchi's model?

Secchi was an established scientist and well recognized throughout the western world, especially in Roman Catholic France. Secchi's first Italian paper in the *Bullettino Meteorologico* had already been published for nearly one year [95] by the time Faye gave his address [111]. Secchi's second paper on the constitution of the photosphere was immediately translated into *Les Mondes* by l'Abbé Moigno. It appeared in Paris on December 22nd, 1864 [96]. This was nearly one month prior to Faye's presentation before *l'Académie des Sciences* on January 16th. Faye's first paper was silent on this point. Nonetheless, in his second paper, presented on January 25th of the same year, Faye reported that "*I have seen, a few days ago, a correspondence by Father Secchi, who has much too studied the Sun to share the popular view reigning today on the liquidity of the photosphere, that our corresponding scientist has arrived from his side to an explanation of sunspots founded on the same principle*[1]" [112, p. 146]. The footnote in Faye's sentence referred to Moigno's translation of Secchi's second paper [96].

Faye's second paper began with a discussion of solar rotation and particularly of the work of Carrington [112, p. 140–142]. He then discussed Helmholtz' contraction hypothesis [112, p. 143] and highlighted the enormous temperatures inside the Sun as a cause of the complete dissociation of its constituents. These gases rose to the solar exterior where they condensed into non-gaseous particles susceptible to incandescence. Faye reasoned that the formation of the photosphere was simply a consequence of the cooling of internal gases [112, p. 144]. He reconciled Arago's argument on polarization with Kirchhoff's need for a continuous spectrum [112, p. 145]. In so doing, he advanced a photosphere based essentially on Secchi's model when he described: *incandescent particles, floating on a gaseous medium*" [111, p. 145]. Faye then highlighted that sunspots were produced by the visualization of the gaseous solar interior [112, p. 146]. This became the source of Spencer's "famous objection" in *The Reader* [105] and reflected Faye's incomplete comprehension of thermal emission.

Faye closed his second paper with an elaborate description of the vertical convection currents which he postulated were present inside the Sun. He replayed much of Spencer's ideas on the Nebular hypothesis and solar cooling. The Frenchman stated that, given sufficient time, the photosphere would become very thick with the "*consistence of a liquid or a paste*". Herein, he directly linked his ideas to Spencer's liquid photosphere [104]. Hence, along with the arguments based on convection currents, Faye introduced another source of priority claims for the British scholar. Faye's initial exposition [111, 112] was more extensive than Secchi's [95, 96], but not significantly superior to Spencer's [104, 105].

Once his papers on the Constitution of the Sun were presented to the Académie, Faye published a slightly different work in *Les Mondes* [143] in which he again stated that Father Secchi arrived at the same conclusion regarding the photosphere. The Frenchman sought Secchi's approbation [143, p. 298]. As for Secchi, he gallantly responded to Faye's *Les Mondes* article in a letter published in *Comptes Rendus*, on March 6th, 1865 [144]. Secchi wrote in most charitable terms, as if delighted by Faye's claim of simultaneous discovery. If anything improper had occurred, it was silently forgiven. A few years later, in 1867, Secchi would receive *la croix d'officier de la Légion d'Honneur* from the hand of Napoleon III [128, p. iii, 208].

Faye first addressed the sunspot problem in his model within his third paper on the constitution of the Sun, published in 1866 [120]. He began the discourse by praising English astronomy and citing every prominent British astronomer of the period, including Herschel, Carrington, Dawes, Nasmyth, Stone, Huggins, de la Rue, Stewart, Thomson, and Waterston. Spencer was absent from the list. Still, the focus of Faye's work was a direct address of Spencer's complaint with respect to solar opacity: "*The difficulty is relative to the explanation of sunspots. We know that gases heated to the point of becoming luminous never rise to the point of incandescence; the latter being a property of solid particles, even when they are reduced to the same tenuousness*" [120]. Faye restated Secchi's idea that the photosphere was made of fine condensed incandescent particles floating in a gaseous medium. If these particles were missing from a region, it would necessarily become obscure. This was his explanation of sunspots: regions devoid of these incandescent particles. Faye then raised the "famous objection", without mentioning Spencer's name, as if the charge had come from nowhere: "*In this we object that if gases emit but little light, by consequence they are transparent. If then an opening was made in the photosphere, one should see, across the gaseous internal mass of the Sun, the opposite region of the same photosphere with a brilliance barely diminished; as a result there would no longer be any spots*" [120]. It was only later, in 1867, that Faye was finally forced to acknowledge Spencer as a source of the complaint [122, p. 404]. He did so in a footnote, while insisting that the reproach had first been brought to his attention by the editor of *Comptes Rendus*. This was the most assured means of preventing impropriety. In the same work, Faye remained silent on Spencer's convection currents, variations in solar density, and justified priority claim for a gaseous solar interior.

Faye addressed the complaint by arguing that, in fact, it was a property of gases or vapors to extinguish light as well as an opaque body, provided that the thickness of the gas was sufficient. Faye was essentially invoking optical thick-

ness and, once again, foreshadowing the modern stellar opacity problem. In answering Secchi [144], Faye presented his idea that the interior of the Sun could be viewed as concentric layers of gas [145, p. 296]. The thought was to remain associated with the treatment of the internal constitution of the Sun and was also used by Eddington in advancing his theoretical treatment of the problem [19].

As for Faye's debate with Kirchhoff, it was less than cordial. The battle began when Faye improperly described Kirchhoff's model in the literature [120]. Kirchhoff would rebuke Faye for maintaining that horizontal convection currents did not occur at the level of the photosphere: "*Mr. Faye then rejects the existence in the solar atmosphere of horizontal currents which, in my hypothesis, must explain the different movements of sunspots*" [121, p. 398]. Unlike Kirchhoff, Faye invoked internal convection currents with a vertical displacement. On the surface of the Sun, he wanted voids to obtain the spots, not horizontal currents [122, p. 403]. Faye responded to the father of spectral analysis in the most inappropriate tone: "*I congratulate myself in having received a personal intervention from Mr. Kirchhoff, because his letter explains to me something of which I have always been profoundly astonished, to know the persistence with which a man of such high merit can sustain a hypothesis so incompatible with the best known facts*" [122, p. 401]. Faye, of course, referred to Kirchhoff's cloud model of sunspots. In any case, Faye's arrogance in the published article was met eventually by a sound defeat at the hand of Kirchhoff [124].

Faye was so concerned by Kirchhoff's first letter of objection that he drafted a second response, which was mathematical in nature [123], even before the German had the opportunity of reply to his first answer [122]. In this letter, the Frenchman invoked that the nature of sunspots was similar to the darkened grid associated with solar granulation. He went on to dispute, like his mentor Arago (see Section 3.3), the existence of the corona [123]. Both statements were erroneous. Then, Faye opened a new line of defense for his sunspot theory and the controversy relative to seeing through the Sun. He believed that he could counter Kirchhoff and Spencer by advancing that the gas density inside the Sun was not homogeneous. He began by arguing that the interior of the Sun was highly variable in density [123, p. 222–223]: "*In consequence this central density must be many hundreds or even thousands of times superior to that of the superficial layer which forms the photosphere*". Once again, he failed to credit Spencer, this time regarding varying internal solar densities [105]. Faye then proposed a gaseous internal medium which could be viewed as spherical layers of material [123, p. 222–223]. He advanced the same idea a year earlier during a discussion with Father Secchi [146]. The concept has remained in astronomy to the present.

Finally, Faye made his critical misstep. He invoked that a ray of light which hit the higher density of the mass inside the Sun was refracted inward and unable to escape. The astronomer then audaciously charged Kirchhoff with failing to understand the consequences of a non-homogeneous solar interior.

Kirchhoff was severe in his defense. Using his law of thermal emission, Kirchhoff disarmed Faye. He reminded the scholar that the radiation inside an opaque enclosure must be black [124]. As such, Kirchhoff was, ironically, the first person to postulate that the radiation inside a gaseous Sun, surrounded by an enclosing photosphere, must be black. In reality, Kirchhoff's conclusion was only partially correct. The solar photosphere produced a thermal spectrum. However, it was not truly black, since the Sun maintained convection currents which prevented this possibility. Nonetheless, if the photosphere was condensed and perfectly enclosed a gaseous solar body, then that interior would have to contain the same thermal radiation as emitted on the solar surface. Still, Kirchhoff was mistaken in believing that the radiation would have to be black. It would take many years before this reality became apparent [61–66]. In any case, Kirchhoff's arguments, though not completely sound, well surpassed Faye's physical knowledge of the problem. With time, the modern theory of the Sun eventually applied Kirchhoff's ideas to the problem of internal stellar opacities. In doing so, it removed the condensed nature of the photosphere as a primary source of photons. Therefore, there was a great difference between the problem addressed by Faye and Kirchhoff and the current gaseous models of the stars. Kirchhoff and Faye were dealing with photons produced initially by condensed matter in the photosphere. The modern theory holds that such photons could be generated in the solar core, without recourse to condensed matter and without having the Sun enclosed by its condensed photosphere.

The great battle between Faye and Kirchhoff over the nature of sunspots and the solar constitution would end with a whimper. Faye advanced [125] that Kirchhoff had abandoned his model, because the German failed to defend it in his rebuttal letter [124]. Kirchhoff retorted by emphatically arguing that he continued to defend his solar theory [126].

As for Faye, he was completely unable to respond to Kirchhoff's closing argument on the presence of blackbody radiation inside a gaseous solar model. In 1872, he finally abandoned his first theory of sunspots, replacing it with cyclonic formation, an idea for which he once again failed to credit Spencer. Yet, in closing the openings he had created in the photosphere, Faye finally referred to Spencer [119] for his "famous objection". By this time, the problem of internal solar opacity had become irrelevant. Mankind became, at least for the moment, theoretically unable to "see within the Sun". The fully gaseous models, advanced in the 20th century, reintroduced the concept that scientists could visualize differing depths within the Sun. Despite the lack of the enclosure, as required by Kirchhoff in his 1867 letter [124], the modern solar interior has been hypothesized to contain blackbody radiation [15–17].

As a point of interest, the differences between Faye's, Secchi's, and Lockyer's concepts of sunspots have been reviewed in the 1896 version of Young's classic text [8, p. 182–190]. Today, nearly all of these ideas have been abandoned. Much of the controversies which called for the dismissal of Kirchhoff's condensed photosphere have long ago evaporated. The Wilson effect alone remains [84], as a standing tribute to that great English astronomer, who unlike Faye and many of his contemporaries, was so careful relative to queries and conjectures.

4.6 Discord, stellar opacity, and the birth of the gaseous Sun

Imagine a gaseous Sun. The idea was so tantalizing for men of the period that it became a source of instant quarrel for priority. Secchi gently rebuked Kirchhoff [95], absolved Faye [144], and defended himself against Lockyer [5, p. 500–515]. Faye, in turn, battled with Kirchhoff [121–127] and after securing the blessing of Father Secchi [144], was quick to announce his innocence before the Académie: "*This letter [from Secchi] demonstrates that we followed at the same time, Father Secchi and I, a train of ideas which was altogether similar...*" [145, p. 468]. Like his English counterparts, Faye acted as if he was also unaware of John Herschel's 1864 article [97]. But what could be said of this coincidence of ideas? Was it really possible that, in the span of a few months, Secchi, Herschel, Faye, Lockyer and Frankland, and de la Rue along with Stewart and Loewy all independently conceived of the same idea? Faye addressed the question: "*With respect to the analogies that Father Secchi signals with reason between his ideas and mine, coincidences of this type offer nothing which can surprise, identical ones [ideas] are produced every time that a question is ripe and is ready for a solution*" [145]. But surely, the argument could not be extended to every prominent astronomer of the period. Being first and very likely ignorant of Spencer's English text [104], only Secchi could claim truly independent thought.

After hearing from the Jesuit astronomer, Faye finally cited Magnus [145, p. 471], the scientific element which was central to his model, but which, unlike Secchi, he had so neglected in his earlier works. However, if one accounted for Spencer's and Secchi's ideas in Faye's famous papers [111, 112], there was not much left as original thought. The most significant exception was Faye's idea that the photosphere of the Sun was devoid of a real surface [13, p. 42], also advanced in *Les Mondes* [143]. Faye believed that the "*presence of the photosphere does not interrupt the continuity of the [central] mass*" of the Sun [143, p. 301] and insisted that "*This limit is in any case only apparent, the general milieu where the photosphere is incessantly forming surpasses without doubt more or less the highest crests or the summits of the incandescent clouds*" [143, p. 298]. Such was the first consequence of the gaseous models: there could be no defined solar surface. The problem continues to haunt astrophysics to this day [57, 146]. With Faye, the Sun lost its distinct surface.

It is evident that Faye never properly acknowledged Spencer [120, p. 235]. Nonetheless, he remained delighted that his works had been immediately reviewed in *The Reader* by Lockyer, as evidenced by his 1865 letter [145]. As such, it is doubtful, as early as 1865, that he never knew of Spencer's rebuttal [105]. Faye behaved as if concerns against his "*transparent solar interior*" originated exclusively from Kirchhoff as late as 1866 [121]. In fact, it was clear that the criticism of seeing through the Sun had been swiftly leveled by Spencer [105, p. 228]. Since Kirchhoff was a friend of Roscoe [61], it was not unlikely that he quickly became aware of *The Reader* series. Once again, Spencer wrote: "*But if these interior gases are non-luminous from the absence of precipitated matter must they not for the same reason be transparent? And if transparent, will not the light from the remote side of the photosphere, seen through them, be nearly as bright as that from the side next to us?*" [105, p. 228]. Meadows argued that this criticism of Faye's work originated from Balfour Stewart [13, p. 41–42], but did so without citation. In fact, the reference to Balfour Stewart was provided by Norman Lockyer, when he reprinted his letters, in 1874, and added a footnote giving credit to Balfour Stewart over Kirchhoff [5, p. 57], well after Spencer made his case. This was how Lockyer distorted the scientific record using a footnote: "*This note was added to the article as it originally appeared, as the result of a conversation with my friend Dr. Balfour Stewart. I am more anxious to state this, as to him belongs the credit of the objection, although, as it was some time afterwards put forward by Kirchhoff, the latter is now credited with it, although it was noticed by Faye, Comptes Rendus, vol. lxiii, p. 235, 1866. The idea is this: — If the interior solar gases are feeble radiators, then, on the theory of exchanges, they must be feeble absorbers; hence they will be incompetent to absorb the light coming through the hypothetically gaseous Sun from the photosphere on the other side (1873)*" [5, p. 57]. One can only wonder why the discoverer of Helium, one of the great fathers of spectral analysis, and the founder of the journal *Nature*, insisted on altering the historical record. Apparently, Spencer was not as weak in thermodynamics, as previously argued [106].

4.7 Stewart, Kirchhoff, and amateurs

Stewart had been an author on the initial paper with de la Rue and Loewy [133–135]. But suddenly, he detached himself from this position when he discussed the photosphere, without invoking the presence of a gas: "*Next with regard to the photosphere or luminous envelope of the Sun, this surface, when viewed through powerful telescopes, appears granulated or mottled... But besides this there is reason to believe that great defining as well as magnifying power discloses the fact that the whole photosphere of the Sun is made up of detached bodies, interlacing one another, and preserving a great amount of regularity both in form and size*" [147]. Thus, when Stewart wrote independently, it was obvious that he ac-

tually believed that the photosphere was a liquid or solid. In this respect, he became aligned with Spencer and Kirchhoff on the condensed nature of the photosphere.

In his *Lessons in Elementary Physics*, Stewart persisted in breaking from de la Rue and Loewy [148, p. 279]. This was the case even in the edition published closest to the end of his life. In this classic text for its day, Stewart stated: "*If we throw upon the slit of our spectroscope an image of the Sun or of one of the stars, with the view of obtaining its spectrum, we find a large number of black or dark lines in a spectrum otherwise continuous, and we argue from this that in the Sun or stars we start with a solid or liquid substance, or at any rate with some substance which gives us a continuous spectrum, and that between this and the eye we have, forming a solar or stellar atmosphere, a layer of gas or vapours of a comparatively low temperature, each of which produces its appropriate spectral lines, only dark on account of the temperature of the vapours being lower than that of the substance which gives the continuous spectrum*" [148, p. 279]. Again, there was no mention of a gaseous photosphere supporting condensed matter precipitates in this description of the problem. In fact, this passage echoed Kirchhoff's explanation [43], as Stewart was all too aware of the nature of thermal emission in gases [149].

Hence, the Scottish physicist very much desired that the photosphere be condensed, as evidenced initially in his 1864 article: *On the Origin of Light in the Sun and Stars* [150]. In this work, Stewart advanced that planets could alter the brightness of stars by modifying the amount of sunspots. He tried to answer the question "*From all this it is evident that in the case of many stars we cannot suppose the light to be due to an incandescent solid or liquid body, otherwise how can we account for their long continued disappearance?*" [150, p. 452]. The entire manuscript was aimed at accounting for this disappearance, even if the photosphere was solid or liquid. He stated in this regard "*if it can be proved, as we think it can, that a disc full of spots is deficient in luminosity*" [150, p. 452]. Stewart made this conjecture to explain the occurrence of variables [150]. For him, the photosphere had to be liquid or solid. But variable stars posed a tremendous scientific difficulty. As a result, he required something like planets to modify their emission cycles [150]. Stewart reconciled his desire for a liquid or solid photosphere within these types of stars by stating: "*the approach of a planet to the Sun is favourable to luminosity*" [150, p. 454]. His desire for condensed matter was so powerful that Stewart advocated the scientific error that Venus itself can modify the appearance of sunspots [150, p. 454]. Regrettably, Stewart would eventually discover Loewy's misconduct while producing mathematical reductions relative to the work at Kew [151, p. 361]. This would place a considerable tarnish on the Kew group, and Stewart would never again speak on planetary effects relative to sunspots.

Earlier, in *Origin of Light* [150, p. 450–451] Stewart had viewed sunspots as cavities on the Sun, produced by an opening in the photospheric matter revealing the dark nucleus of the interior. In 1864, just prior to the paper with de la Rue and Loewy, Stewart stated that the Sun possessed with a solid body [150, p. 451]. The concept was similar to Wilson [84].

Despite Loewy's misconduct [151], Stewart could not long maintain a fully gaseous photosphere, given his extensive knowledge of thermal emission in gases [149]. Clearly, he had not embraced de la Rue's model [133–135] and the claim by Lockyer, discussed in Section 4.7, that the photosphere could be completely gaseous and devoid of any condensed matter [37]. On the same note, Stewart's entire discussion on thermal radiation, in his classic physics text, is well worth reading [148, p. 270–297]. It revealed his profound knowledge of such processes and also his understanding that gases cannot produce the continuous spectrum required.

Stewart maintained support for what is essentially Kirchhoff's liquid photospheric model. He did so despite his previous adversity with the German [61, 139]. In this regard, he was being guided by the same scientific reasoning as his former detractor [43]. The Scottish scientist also held profound values [140, 141, 150]. As such, it is comforting to notice how, in some sense, the two men were now reconciled. Stewart's continued support for Kirchhoff's condensed photosphere, was astounding as it *de facto* dismissed any previous arguments relative to Andrew's critical temperature [28] and line broadening [37]. For Stewart, the primary determinant of the phase of the photosphere was its thermal emission. The same held true for Kirchhoff. Yet, Stewart's insistence was important because it continued well after critical temperatures and line broadening had entered the halls of astronomy. Those who maintained that the photosphere was gaseous, therefore, continued alone on their journey. They marched on without the support of the two great experts in thermal radiation: Gustav Kirchhoff and Balfour Stewart.

As for Spencer, if there was any merit in his work, other than his obvious and justified claim of priority, it was that he foresaw internal convection currents, variable solar density, and the tremendous problem of internal stellar opacity. The last of these, contained in the "famous objection", remains a key problem with the idea of a gaseous Sun, despite all attempts to rectify the situation [69, 70]. But what is most fascinating about this philosopher, remains his amateur status in astronomy. Karl Hufbauer has commented on the contributions of amateurs to astrophysics [152]. Bartholomew argues as though there was little room for Spencer and his theoretical ideas in solar science [106]. In this regard, he stands in profound opposition to George Hale, one of the greatest solar observers and the founder of the *Astrophysical Journal*. In 1913, Hale defended the special place of amateurs in astronomy when he drafted the moving obituary of Sir William Huggins: "*If it be true that modern observatories, with their expensive equipment, tend to discourage the serious amateur, then it may be doubted whether the best use is being made of the funds they represent. For the history of sci-*

ence teaches that original ideas and new methods, as well as great discoveries resulting from the patient accumulation of observations, frequently come from the amateur. To hinder his work in any serious way might conceivably do a greater injury than a large observatory could make good... Every investigator may find useful and inspiring suggestions in the life and example of Sir William Huggins. Their surest message and strongest appeal will be to the amateur with limited instrumental means, and to the man, however situated, who would break new ground" [153].

Notes and acknowledgements

The author would like to thank Professor Eileen Reeves (Department of Comparative Literature, Princeton), along with Mary Posani (Department of French and Italian, The Ohio State University), for their translations of Secchi's key Italian papers on the solar constitution. All translations from French were generated by the author. He would also like to acknowledge the efforts of Bernadette Carstensen (Circleville, Ohio) for the graphite illustrations of Spencer, Secchi, and Faye. The author recognizes the extensive assistance he received over the years from Rebecca Jewett, Assistant Curator for Rare Books and Manuscripts, along with Mary Reis (Microforms Section), from the Thomson Library of The Ohio State University. The numerous articles appearing in the *Comptes Rendus hebdomadaires des séances de l'Académie des sciences* can be accessed online, without charge, through the French National Library (http://gallica.bnf.fr). Several of the older texts can be found online using either books.google.com or other digital sources. The articles published in *Les Mondes*, along with several key texts, were accessed by purchasing the requisite volumes.

Dedication

This work is dedicated to my youngest son, Luc.

Submitted on April 16, 2011 / Accepted on April 23, 2011
First published in online on May 07, 2011

References

1. Grant R.G. History of Physical Astronomy from the Earliest Ages to the Middle of the Nineteenth Century. Henry G. Bohn, London, 1852.
2. Lewis G.C. An Historical Survey of the Astronomy of the Ancients. Parker, son, and Bourn, West Strand, London, 1862.
3. Secchi A. Le Soleil, Gauthier-Villars, Paris, 1870.
4. Proctor R.A. The Sun: Ruler, Fire, Light, and Life of the Planetary System. Longmans, Green, and Co., London, 1872.
5. Lockyer J.N. Contributions to Solar Physics. MacMillan and Co., London, 1874.
6. Newcomb S. Popular Astronomy. Harper & Brothers, New York, 1882.
7. Clerke A.M. A Popular History of Astronomy during the Nineteenth Century. 3rd Edition, Adam and Charles Black, London, 1893.
8. Young C.A. The Sun and the Phenomena of Its Atmosphere. Charles C. Chatfield & Co., New Haven, 1872; also a latter edition: The Sun. D. Appleton & Co., New York, 1896.
9. Fison A.H. Recent Advances in Astronomy. Blackie and Son, London, 1898. 1–50.
10. Berry A. A Short History of Astronomy. Charles Scribner's Sons, New York, 1899.
11. Clerke A.C. The System of the Stars. Longmans, Green, and Co., London, 1890.
12. Maunder E.W. Are the Planets Inhabited? Harper & Brothers, London, 1913.
13. Meadows A.J. Early Solar Physics. Pergamon Press, Oxford, 1970.
14. Hufbauer K. Exploring the Sun: Solar Science since Galileo. John Hopkins University Press, Baltimore, 1991.
15. Kippenhahn R. and Weigert A. Stellar Structure and Evolution. Springer-Verlag. Berlin, 3rd printing, 1994.
16. Clayton D.D. Principles of Stellar Evolution and Nucleosynthesis. McGraw-Hill, New York, 1968.
17. Reddish V.C. The Physics of Stellar Interiors: An introduction. Edinburgh University Press, Edinburgh, 1974.
18. Sturrock P.A., Holzer T.E., Mihalas D.M. and Ulrich R.K. Physics of the Sun (Volume 1: The Solar Interior). D. Reidel Publishing Co., Dordrecht, 1986.
19. Eddington A.S. The Internal Constitution of the Stars. Dover Publ. Inc., New York, 1959.
20. Eddington A.S. Stars and Atoms. Yale Univ.Press, New Haven, 1927
21. Chandrasekhar S. An Introduction to the Study of Stellar Structure. Dover Publ., New York, 1939.
22. Laplace P.S. Exposition du système du monde. Imprimerie du Cercle-Social, Paris, 1796 (available online: http://dx.doi.org/10.3931/e-rara-497; Also available in English: Pond J. The system of the world, London, 1809).
23. Numbers R.L. Creation by Natural Law: Laplace's Nebular Hypothesis in American Thought. Seattle, 1977, 124–32.
24. von Helmholtz H. On the interaction of natural forces. *Phil. Mag.*, 4th series, 1856, v.11, 489–517.
25. von Helmholtz H. "Über die wechselwirkungen der naturkräfte" — On the interaction of natural forces: A lecture delivered February 7, 1854 at Königsberg in Prussia. In *Popular Lectures on Scientific Subjects*, D. Appleton and Company, New York, 1873, 153–196.
26. Cagniard de la Tour C. Exposé de quelques résultats obtenu par l'action combinée de la chaleur et de la compression sur certains liquides, tels que l'eau, l'alcool, l'éther sulfurique et l'essence de pétrole rectifiée, *Ann. Chim. Phys.*, 1822, v.21, 127–132; Supplément, ibid., 178–182.
27. Berche B., Henkel M., and Kenna R. Critical phenomena: 150 years since Cagniard de la Tour. *J. Phys. Studies*, 2009, v.13(3), 3201/1–3201/4.
28. Andrews T. The Bakerian lecture: On the continuity of the gaseous and liquid states of matter. *Phil. Trans. Roy. Soc.*, 1869, v.159, 575–590.
29. Rowlinson J.S. Thomas Andrews and the critical point. *Nature*, 1969, v.224, 541–543.
30. 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.
31. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsverögen und dem Absorptionsvermögen. der Körper für 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.*, ser. 4, 1860, v.20, 1–21.)
32. Kirchhoff G. On the relation between the emissive and the absorptive power of bodies for light and heat. (Reprinted from: *Investigations of*

the Solar Spectrum and the Spectra of the Chemical Elements, 2nd edition, Berlin, Ferd. Dummler's Publishing House, 1866, Gesammelte Abhandlungen, 571–598, Leipzig, 1882 as found in Harper's Scientific Memoirs (J.S. Ames, Ed.) — 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.

33. Plücker J. and Hittorf J.W. On the spectra of ignited gases and vapours, with especial regard to the different spectra of the same elementary gaseous substance. *Phil. Trans. Roy. Soc.*, London, 1865, v.155, 1–29.
34. Wüllner M.A. On the spectra of certain gases in Geissler's tubes. *Phil. Mag.*, 1869, v.37, 405–424 (translated from original: Wüllner M.A. Ueber die Spectra einiger Gase in Geifsler'schen Röhren. *Poggendorff's Annalen der Physik und Chemie*, 1868, v.135, 497–527).
35. Wüllner M.A. On the spectra of some gases under high pressure. *Phil. Mag.*, 1870, v.39, 365–370 (translated from original: Wüllner M.A. Ueber die Spectra einiger Gase bei hohem Drucke. *Poggendorff's Annalen der Physik und Chemie*, 1869, v.137, 337–361).
36. Frankland E. On the combustion of hydrogen and carbonic acid in oxygen under great pressure. *Proc. R. Soc. Lond.*, 1867, v.16, 419–422.
37. Frankland E. and Lockyer J.N. Preliminary note of researches on gaseous spectra in relation to the physical constitution of the Sun. *Proc. Roy. Soc. London*, 1869, v.17, 288–91 (also found in ref. 5, 525–529 and in ref. 13, 203–206).
38. Swedenborg E. Opera Philosophica et Mineralia. 1734.
39. Swedenborg E. The Principia; or the first principles of natural things, being new attempts towards a philosophical explanation of the elementary world. Translated by: Augustus Clissold, W. Newbery, London, 1846.
40. Whitaker D.K., Clapp M., Simms W.G., and Thornwell J.H. The philosophical character of Swedenborg. *Southern Quarterly Review*, 1848, v.13(26), 427–469.
41. Woolson M.M. Solar system-its origin and evolution. *Quarterly J. Roy. Astron. Soc.*, 1993, v.34, 1–20.
42. Woolson M.M. Solar system-its origin and evolution: A personal view. *Quarterly J. Roy. Astron. Soc.*, 1993, v.34, 101–102.
43. Kirchhoff G. The physical constitution of the Sun. In *Researches on the solar spectrum and the spectra of the chemical elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, 23–31.
44. Mayer J.R. Die Entstehung der Sonnenwärme. in *Beitraege zur Dynamic des Himmels in populärer Darstellung*, Verlag von Johann Ulrich Landherr, Heilbronn, 1848, 20–28.
45. Mayer J.R. Bemerkungen über die Kräfte der undelebten Natur. *Annalen der Chemie und Pharmacie*, 1842, v.42, 233–240 (reprinted in Mayer J.R. Die Mechanik der Wärme: in gesammelten Schriften, Verlag der J.G. Gotta'schen Buchhandlung, Stuttgart, 1867, 1–12).
46. Thomson W. On the mechanical energies of the solar system. *Trans. Royal Soc. Edinburgh*, April, 1854, v.21, 64–80; and *Phil. Mag.*, 4th series, 1854, v.8, 409–430.
47. Thomson W. On the age of the Sun's heat. *Macmillan's Magazine*, 1862, v.5, 288–293.
48. Boyle R. New Experiments Physico-Mechanical, Touching the Spring of the Air, and Its Effects (Made, for the Most Part in a New Pneumatical Engine) Written by Way of Letter to the Right Honorable Charles Lord Vicount of Dungarvan, Eldest Son to the Earl of Corke. H. Hall, Oxford, 1660.
49. Gay-Lussac J.L. Recherches sur la dilatation des gaz et des vapeur. *Annales de Chimie*, 1802, v.43, 137–175.
50. Clapeyron E. Mémoire sur la puissance motrice de la chaleur. *Journal de l'École Polytechnique*, 1834, v.14, 153–190.
51. Eddington A.S. The internal constitution of the stars. *The Observatory*, 1920, v.43 (557), 341–357.
52. Bethe H.A. Energy production in stars. *Phys. Rev.*, 1939, v.55, 434–56.
53. Haynes W.M. CRC Handbook of Chemistry and Physics. 91st Edition, CRC Press, 2010–2011, 4–121.
54. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v.3, 764–770.
55. Weir S.T., Mitchell A.C. and Nellis W.J. Metallization of fluid molecular hydrogen at 140 GPa (140 Mbar). *Phys. Rev. Letters*, 1996, v.76, 1860–1863.
56. Babaev E, Sudbø A. and Ashcroft N.W. A superconductor to superfluid phase transition in liquid metallic hydrogen. *Nature*, 2004, v.431, 666–668.
57. Robitaille P.M. The solar photosphere: evidence for condensed matter. *Progr. Phys.*, 2006, v.2, 17–21.
58. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v.1, 70–81; arXiv: astro-ph/0410075.
59. Robitaille P.M. The Sun as a high energy/high density liquid metallic hydrogen plasma, *The 33rd IEEE International Conference on Plasma Science*, June 4–8, 2006, Traverse City, Michigan, p. 461, DOI: 10.1109/PLASMA.2006.1707334.
60. Robitaille P.M. A radically different point of view on the CMB. In: *Questions of Modern Cosmology — Galileo's Legacy*, ed. by M. D'Onofrio and C. Burigana, Springer, New York, 2009.
61. Robitaille P.M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v.3, 36–55.
62. Robitaille P.M. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v.31(6), 1263–1267.
63. Robitaille P.M. An analysis of universality in blackbody radiation. *Progr. Phys.*, 2006, v.2, 22–23; arXiv: physics/0507007.
64. Robitaille P.M. A critical analysis of universality and Kirchhoff's law: a return to Stewart's law of thermal emission. *Progr. Phys.*, 2008, v.3, 30–35; arXiv: 0805.1625.
65. Robitaille P.M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v.4, 3–13.
66. Robitaille P.M. The little heat engine: Heat transfer in solids, liquids, and gases. *Progr. Phys.*, 2007, v.4, 25–33.
67. Planck M. Über eine Verbesserung der Wien'schen Spectralgleichung. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 1900, v.2, 202–204. (This is Planck's famous October 19, 1900 lecture. It can also 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, 3–5 or 35–37.)
68. 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. In: *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.)
69. Planck M. The Theory of Heat Radiation. P. Blakiston's Son & Co., Philadelphia, PA, 1914.
70. 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(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).

71. Langley S.P. Experimental determination of wave-lengths in the invisible spectrum. *Mem. Natl. Acad. Sci.*, 1883, v.2, 147–162.
72. Langley S.P. On hitherto unrecognized wave-lengths. *Phil. Mag.*, 1886, v.22, 149–173.
73. Langley S.P. The invisible solar and lunar spectrum. *Am. J. Science*, 1888, v.36(216), 397–410.
74. The Opacity Project Team. The Opacity Project. Institute of Physics Publishing, Bristol, UK, 1995, v.1.
75. The Opacity Project Team. The Opacity Project. Institute of Physics Publishing, Bristol, UK, 1996, v.2.
76. Hansen C.J., Kawaler S.D. and Trimble V. Stellar Interiors: Physical Principles, Structure, and Evolution. Springer-Verlag, New York, 2004.
77. Humphreys W.J. Changes in the wave-frequencies of the line of emission spectra of the elements, their dependence upon the elements themselves and upon the physical conditions under which they are produced. *Astrophys. J.*, 1897, v.6(3), 169–232.
78. Lane J.H. On the theoretical temperature of the Sun; under the hypothesis of a gaseous mass maintaining its volume by it internal heat, and depending on the laws of gases as known to terrestrial experiment. *American Journal of Science and Arts*, 1820, July 1870, v.50(148), 57–74.
79. Schuster A. The solar atmosphere. *Astrophys. J.*, 1902, v.16, 320–327.
80. Schuster A. Radiation through a foggy atmosphere. *Astrophys. J.*, 1905, v.21, 1–22.
81. Very F.W. The absorption of radiation by the solar atmosphere and the intrinsic radiation of that atmosphere. *Astrophys. J.*, 1904, v.19, 139–150.
82. Schwarzschild K. Concerning the equilibrium of the solar atmosphere. *Nachrichten von der Koniglichen Gesellschaft der Wissenschaften zu Gottigen*, *Göttinger Nachrichten*, 1906, 195, 41–53 (also found in: Menzel D.H. Selected Papers on the Transfer of Radiation. Dover, New York, 1966; also in: A.J. Meadows [13], 277–290).
83. Herschel W. On the nature and construction of the Sun and fixed stars. *Phil. Trans. Roy. Soc.*, 1795, v.85, 46–72.
84. Wilson A. Observations on the solar spots. *Phil. Trans. Roy. Soc.*, 1774, v.64, 1–30.
85. Herschel W. Experiments on the refrangibility of the invisible rays of the Sun. *Phil. Trans. Roy. Soc.*, 1800, v.90, 284–292.
86. Herschel W. Experiments on the solar, and on the terrestrial rays that occasion heat; with a comparative view of the laws to which light and heat, or rather the rays which occasion them, are subject, in order to determine whether they are the same, or different. Part I. *Phil. Trans. Roy. Soc.*, 1800, v.90, 293–326.
87. Herschel W. Experiments on the solar, and on the terrestrial rays that occasion heat; with a comparative view of the laws to which light and heat, or rather the rays which occasion them, are subject, in order to determine whether they are the same, or different. Part II. *Phil. Trans. Roy. Soc.*, 1800, v.90, 437–538.
88. Herschel W. Observations tending to investigate the nature of the Sun, in order to find the causes or symptoms of its variable emission of light and heat; with remarks on the use that may possibly be drawn from solar observations. *Phil. Trans. Roy. Soc.*, 1801, v.91, 265–318.
89. Arago M. Popular Lectures on Astronomy Delivered at the Royal Observatory of Paris (with extensive additions and corrections by D. Lardner). Greeley and McElrath, New York, 1848.
90. Arago M. Popular Astronomy. Translated by W.H. Smyth and R. Grant, Longman and Co., London, 1855.
91. Herschel J.F.W. Results of Astronomical Observations Made During the Years 1834, 5, 6, 7, 8, at the Cape of Good Hope. Smith, Elder and Co., Cornhill, 1847.
92. Milne E.A. Bakerian lecture: The structure and opacity of a stellar atmosphere. *Phil. Trans. Roy. Soc.*, London 1929, v.228, 421–461.
93. Herschel J.F.W. Outlines of Astronomy. Longman and Co., London, 1849.
94. Herschel J.F.W. Outlines of Astronomy. 10th edition. Longman, Green, and Co., London, 1893.
95. Secchi A. Sulla Teoria Delle Macchie Solari: Proposta dal sig. Kirchoff. *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 31 January 1864, v.3(4), 1–4 (translated into English by Eileen Reeves and Mary Posani: On the Theory of Sunspots Proposed by Signor Kirchoff, *Progr. Phys.*, 2011, v.3, 26–29 — a paper published in this Special Issue).
96. Secchi A. Sulla Struttura della Fotosfera Solare. *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 30 November 1864, v.3(11), 1–3. (translated into French by François Moigno: Sur la structure de la photosphère du soleil. *Les Mondes*, 22 December 1864, v.6, 703–707; translated into English by Mary Posani and Eileen Reeves: On the structure of the photosphere of the sun, *Progr. Phys.*, 2011, v.3, 30–32 — a paper published in this Special Issue).
97. Herschel J. On the solar spots. *Quarterly Journal of Science*, 1864, v.1, 219–235.
98. Lalande J. Abrégé d'astronomie. Barthelemi Vlam, Amsterdam, 1774.
99. Descartes R. Principia Philosophiae. Apud Ludovicum Elzevirium, Amstelodami, 1644.
100. Descartes R. Principles of Philosophy. Translated by V.R. Miller and R.P. Miller, D. Reidel, Dordrecht, 1983.
101. Galilei G. and Scheiner C. On Sunspots. Translated by E. Reeves and A. V. Helden, University of Chicago Press, Chicago, 2010.
102. Magnus G. Notice sur la constitution du soleil. *Archives des science physique et naturelles (Genève)*, 1864, v.20, 171–175; (translated from: Magnus G. Notiz über die Beschaffenheit der Sonne. *Poggendorff's Annalen der Physik und Chemie*, 1864, v.121, 510–512; translated into English by Patrice Robitaille: Notice on the Constitution of the Sun. *Progr. Phys.*, 2011, v.3 — a paper published in this Special Issue).
103. Lovering J. Address before the American Association. *The Popular Science Monthly*, December 1874, 197–212.
104. Unsigned (Spencer H.) Recent astronomy, and the nebular hypothesis. *Westminster Review*, 1858, v.70, 185–225.
105. Spencer H. The constitution of the sun. *The Reader: A Review of Current Literature*, 25 February 1865, v.5, 227–229 (also found in: Spencer H. Essays: Scientific, Political, and Speculative. Vol. III. Williams and Norgate, London, 1875, 217–229).
106. Bartholomew C.F. Herbert Spencer's contributions to solar physics. *J. Hist. Astronomy*, 1988, v.19, 1–28
107. Spencer H. The Principles of Biology. Volume 2. Williams and Norgate, London, 1867.
108. Youmans E.L. (Editor) Spencer-smashing in Washington. *Popular Science Monthly*, 1894, v.44 , 856–857.
109. Lockyer N. On the physical constitution of the Sun. *The Reader: A Review of Current Literature*, 28 January 1865, v.5, 107 (see also [5], 44–50).
110. Lockyer N. On the physical constitution of the Sun (second article). *The Reader: A Review of Current Literature*, 4 February 1865, v.5, 140–142 (see also [5], 51–62).
111. Faye H.A.E.A. Sur la constitution physique du Soleil — première partie. *Comptes Rendus*, 1865, v.60, 89–96.
112. Faye H.A.E.A. Sur la constitution physique du Soleil – deuxième partie. *Comptes Rendus*, 1865, v.60, 138–150.
113. Yeo R. Science and intellectual authority in mid-nineteenth-century Britain: Robert Chambers and *Vestiges of the Natural History of Creation*, Victorian Studies, 1984, v.28, 5–31.

114. Secchi A. Sur la profondeur des taches et la réfraction de l'atmosphère du Soleil. *Comptes Rendus*, 1866, v.63, 163–170.

115. Secchi A. Résultats de quelques observations spectroscopiques des bords du Soleil — Remarques sur l'obscurité relative des taches solaires — Apparition des étoiles filantes de novembre. *Comptes Rendus*, 1868, v.67, 1018–1022.

116. Secchi A. Remarques sur la relation entre les protubérance et les taches solaires. *Comptes Rendus*, 1869, v.68, 237–238.

117. Secchi A. Resultats fournis par l'analyse spectrale de la lumière d'Uranus, de l'étoile R des Gémeaus, et des taches solaires. *Comptes Rendus*, 1869, v.68, 761–765.

118. Secchi A. Sur les divers aspects des protubérances et des autres parties remarquables, à la surface du Soleil. Classification de ces phénomènes. *Comptes Rendus*, 1871, v.73, 826–836.

119. Faye H.A.E.A. Completement de la théorie physique du Soleil; explication des taches. *Comptes Rendus*, 1872, v.75, 1664–1672.

120. Faye H. Sur quelques objections relatives à la constitution physique du Soleil. *Comptes Rendus*, 1866, v.63, 234–237.

121. Kirchhoff G. Sur les taches solaires. *Comptes Rendus*, 1867, v.64, 396–400.

122. Faye H. Remarques sur la lettre de M. Kirchhoff et sur l'hypothese des nuages solaires. *Comptes Rendus*, 1867, v.64, 400–407.

123. Faye H. La cause et l'explication du phénomène des taches doivent-elles être cherchées en dehors de las surface visible du Soleil? *Comptes Rendus*, 1867, v.65, 221–229.

124. Kirchhoff G. Sur les taches solaires. *Comptes Rendus*, 1867, v.65, 644–646.

125. Faye H. Simple remarque sur la dernière lettre de M. Kirchhoff. *Comptes Rendus*, 1867, v.65, 661–662.

126. Kirchhoff G. Taches solaires: Réponse aux dernière remarques de M. Faye. *Comptes Rendus*, 1867, v.65, 1046.

127. Faye H. Sur l'origine du monde: Théories cosmogoniques des anciens et des modernes. Gauthier-Villars, Paris, 1885.

128. Moigno M. Le Révérend Père Secchi, sa vie, son observatoire, ses travaux, ses écrits, ses titres à la gloire, hommages rendus à sa mémoire, ses grands ouvrages. Librairies des Mondes, Gauthier-Villars, Paris, 1879.

129. Robitaille P.M. Max Karl Ernst Lugwig Planck. *Progr. Phys.*, 2007, v.4, 117–120.

130. Rabounski D. Declaration of academic freedom: Scientific Human Rights. *Prog. Phys.*, 2006, v.1, 57–60.

131. Wallace L., Bernath P., Livingston W., Hinkle K., Busler J., Guo B. and Zhang K.-Q. Water on the sun. *Science*, 1995, v.268, 1155–1158.

132. Carrington R.C. Observations on the Spots of the Sun, from November, 9, 1853, to March 24, 1861, Made at Redhill. Williams and Norgate, London, 1863.

133. de la Rue W., Stewart B. and Loewy B. Researches on solar physics — Series I: On the nature of solar spots (abstract). *Proc. Roy. Soc. London*, 1865, v.14, 37–39.

134. de la Rue W., Stewart B. and Loewy B. Researches on solar physics — Series II: On the behaviour of sun-spots with regard to increase and diminution (abstract). *Proc. Roy. Soc. London*, 1865, v.14, 59–63.

135. de la Rue W., Stewart B. and Loewy B. Researches on Solar Physics (in 4 volumes). Taylor and Francis, London, 1865.

136. de la Rue W., Stewart B. and Loewy B. Researches on solar physics. Heliographical positions and areas of sun-spots observed with the Kew photoheliograph during the years 1862 and 1863. *Phil. Trans. Roy. Soc. London*, 1869, v.159, 1–110.

137. de la Rue W., Stewart B. and Loewy B. Researches on solar physics. No. II. The positions and areas of the spots observed at Kew during the years 1864, 1865, 1866, also the spotted area of the Sun's visible disk from the commencement of 1832 up to may 1868. *Phil. Trans. Roy. Soc. London*, 1870, v.160, 389–496.

138. Clark S. The Sun Kings: The Unexpected Tragedy of Richard Carrington and the Tale of How Modern Astronomy Began. Princeton University Press, Princeton, 2007.

139. Siegel D.M. Balfour Stewart and Gustav Robert Kirchhoff: two independent approaches to Kirchhoff's Law. *Isis*, 1976, v.67(4), 565–600.

140. Schuster A. Memoir of the late Professor Balfour Stewart. *Memoirs and Proceedings of the Manchester Literary and Philosophical Society*, 1888, 4th Series, v.I, 253–272 (contains a full listing of Balfour Stewart's papers).

141. Stewart B. and Tait P.G. The Unseen Universe or, Physical Speculations on a Future State. MacMillan and Co., London, 1875.

142. Godier S., Rozelot J. P. The Solar oblateness and its relationship with the structure of the tacholine and the Sun's subsurface. *Astron. Astrophys.*, 2000, v.355, 365–374.

143. Faye H. Sur la constitution physique du Soleil. Les Mondes, 1865, v.7, 293–306 (translated into English by Patrice Robitaille: On the Physical Constitution of the Sun. *Progr. Phys.*, 2011, v.3, — a paper published in this Special Issue).

144. Secchi A. Lettre à M. Faye sur la constitution du Soleil. *Comptes Rendus*, 1865, v.60, 466–468.

145. Faye H. Remarques sur la lettre du P. Secchi et sur les recherches récemment présentés à la Société Royale de Londres au sujet de la constitution physique du Soleil. *Comptes Rendus*, 1865, v.60, 468–475.

146. Faye H. Sur la réfraction solaire et le dernier Mémoire du P. Secchi. *Comptes Rendus*, 1866, v.63, 193–196.

147. Stewart B. Mr. Balfour Stewart on the latest discoveries concerning the Sun's surface. *Astronomical Register*, 1865, v.29(May), 129–132.

148. Stewart B. Lessons in Elementary Physics. MacMillan and Co., London, 1884.

149. Stewart B. On the radiation and absorption of gases. *Phil. Mag.*, 1863, v.26(174), 219–222.

150. Stewart B. On the origin of the light of the Sun and the stars. *Intellectual Observer*, 1864, v.V., 448–455.

151. Charbonneau P. The rise and fall of the first solar cycle model. *J. Hist. Astron.*, 2002, v.33, 351–372.

152. Hufbauer K. Amateurs and the rise of astrophysics 1840–1910. *Berichte zur Wissenschaftsgeschichte*, 1986, v.9(3), 183–190.

153. Hale G.E. The work of Sir William Huggins. *Astrophys. J.*, 1913, v.37(3), 145–153.

On the Theory of Sunspots Proposed by Signor Kirchoff

A. Secchi

Observatory of the Roman College, Rome, Italy

Eileen Reeves (Department of Comparative Literature, Princeton University, Princeton, New Jersey, 08544) and Mary Posani (Department of French and Italian, The Ohio State University, Columbus, Ohio, 43221) provide a translation of Father Pietro Angelo Secchi's classic work "*Secchi A. Sulla Teoria Delle Macchie Solari: Proposta dal sig. Kirchoff*" as it appeared in *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 31 January 1864, v.3(4), 1–4. This was the first treatise to propose a particulate photosphere floating on the gaseous body of the Sun. The idea would dominate astrophysical thought for the next 50 years. Secchi appears to have drafted the article, as a response to Gustav Kirchhoff's proposal, echoing early Galilean ideas, that sunspots represented clouds which floated above the photosphere. Other than presenting a new solar model, noteworthy aspects of this work include Secchi's appropriate insistence that materials do not emit the same light at the same temperature and his gentle rebuke of Kirchhoff relative to commenting on questions of astronomy.

We gestured in passing in the second number of volume II [of the *Bullettino Meteorologico dell'Osservatorio del Collegio Romano*] to the theory offered by Signor Kirchoff, as a substitution for the current view, about sunspots. This theory has been something of a sensation, since it is the view of a scientist who has rightly gained immense popularity and esteem for his magnificent discoveries concerning the solar spectrum. For this reason, some consideration of his theory is in order, and we will avail ourselves of the various studies that have recently appeared.

Signor Kirchoff rejects both the theory of Herschel and that of Wilson. We will first permit ourselves the observation that it is one thing to refute Herschel's theory, and quite another to refute Wilson's, and that when the first is laid to rest, the second one hardly collapses. Herschel maintained that the solar nucleus was solid, dark, and covered by two layers of luminous clouds, one a certain distance above the other, separated from each other by a non-luminous layer, and he attributed the sunspots to ruptures in these layers. The nuclei formed the body of the sun, which was relatively darker, and visible through the openings in both of these atmospheres; the penumbras were caused, according to Herschel, by the larger rupture in the second luminous layer. Signor Kirchoff does not like the idea of these two atmospheres, and in truth, we have never accepted them either, because they were not necessary, and they were always obliged to rupture together. As a result of our numerous studies, carried out with powerful instruments and with close attention, we concluded that the penumbra was for the most part formed by filamentous currents of the single photosphere that enveloped the sun, or of the same material, rendered so thin that it was transparent. We called attention to the presence of hazes and cirri, lighter than the nuclei, but darker than the penumbras, that were sometimes found within the sunspots; in this we confirmed the discovery of Signor Dawes, who has justifiably complained that until now, no one who studies this phenomena has paid attention to this matter.

Among the issues that have most recently engaged the attention of solar observers is the structure that Signor Nasmyth has called the "willow-leaf" shape. That is, when one observes the sun using reflectors of great size and oculars without darkened lenses,* but in which the light has been weakened, in order to render it tolerable to the eye, by the reflection of a strip of glass, the structure of the sun looks as if it is formed of many elliptical and luminous pieces, elongated in the shape of leaves, and piled one upon the other. They appear more isolated and detached from each other around the penumbras, where they resemble numerous leaves crossing each other, and they are extended in more isolated fashion within the very core of the nucleus.

We have not yet had the opportunity to observe this [willow-leaf] pattern, but we see that even Signor Dawes is in the same circumstances: he finds that the solar structure described by Sir John Herschel, that is, composed of a sort of luminous flakes, is what most closely resembles the appearances observed over the course of many years of research, and in regard to the penumbras, he agrees that there are bright parts, like currents that make their way into the nuclei crossing through the penumbra and retaining all the splendor of the photosphere, and not of the penumbra. This squares with what we ourselves have always observed, and we likewise have always insisted on the three types of substances that are to be seen in each spot: the true nucleus, the penumbra, and the semi-luminous cirri. In order to explain these phenomena, there is no need to rely on two strata of luminous clouds. What suffices, instead, is a simple incandescent photosphere, mixed with less luminous vapors — as one sees in eclipses

**Offuscanti* refer to the dark colored lenses of the type Christoph Scheiner and others put on telescopes if they were observing rather than projecting sunspots.

— in which the ruptures develop, for reasons difficult to ascertain but easy to conjecture, and through which tears one could see the central and less bright part of the star.

But it is precisely this central and darker part that appears a great absurdity to Signor Kirchoff. He asks how it can be maintained that upon contact with such an incandescent body, and under radiation as strong as that of the photosphere, the nucleus itself has not also reached incandescence and fusion. That is [in his view] an absurdity contrary to all the laws of physics. With all the respect that is due to such a distinguished scientist, we believe that this is an exaggeration. First of all, no one has ever said that the nucleus was cold, and if it is dark, it is only in relative terms; Galileo himself said as much in his own epoch, and photography proves the chemical intensity of the nuclei [of sunspots] is so active that in order to obtain an image, one must act instantly, for otherwise the nuclei also are indistinguishable from the photosphere. The difference, therefore, has little to do with their luminosity, and if we were to see one of these nuclei in isolation, perhaps we would hardly be able to distinguish it from an adjacent portion of the sun. Kirchoff relies greatly on the principle that all substances become luminous at the same temperature in order to prove that the core of the sun must be as bright as the photosphere. Here it seems to us that two quite different matters have been conflated: that is, the point at which bodies begin to excite luminous waves capable of being perceptible to the eye, and the fact that all [substances] at the same temperature should be *equally luminous*. We can accept the first of these propositions, and wholly reject the second. In furnaces we see gases of entirely different luminosity from that of solids, and the strongest [hottest] flame that is known — that is, that of the oxyhydrogen blowpipe — is it not one of the least luminous? Thus the conclusion that the parts that form the solar nucleus should be as luminous as the photosphere can hardly be maintained. Nor does it follow that what we call "nucleus" should be either *solid*, or notably less elevated in temperature, but only in a less luminous state; it could even be liquid or gaseous, and only in this state will those lively specific actions that take place in the photosphere fail to occur. The analogy with all planets, as Soret has rightly observed, tells us that the heavier parts should accumulate on the lower stratum, and the lighter ones on the surface, and between these are the gases and the more tenuous materials from whose modifications sunlight is produced. Thus there no longer remains the much-sung absurdity of admitting that beneath the extraordinarily incandescent layer of the photosphere there could be another stratum, perhaps equally warm, but less luminous than it, and that makes itself visible to us when the more incandescent layer of the photosphere itself ruptures.

Moreover, if we reflect carefully, it is not possible to concede an absolute identity in temperature in the various parts of the sun. Indeed, the continuous labor that takes place in that body and the continuous emission of heat suppose that one part must remain in an ongoing state of chemical alteration, and another must be on the verge of entering it; the former might be the photospheric part, and the latter the central and less luminous region, precisely as we observe in ordinary fires. And we would not like to omit the fact that if we were to concede the argument of someone in favor of a sun where all parts are of an equal temperature, that the same could be concluded, following the same logic, about our own furnaces. We are not saying this as if the sun were actually a furnace in which wood were burned; we are saying, rather, that the work itself that takes place to conserve solar activity supposes the existence of some parts that are more intense, and others that are less so. Were this not the case, we would risk regarding the sun as a merely incandescent body, which Thomson has demonstrated could not remain luminous for even a few thousand years.

Treating Wilson's theory as absurd shows that this notion has been confused with that of Herschel, when in fact there is some difference between the two. Wilson said only that the sunspots were cavities, and subsequent observations have verified this *fact*. But no one ever said that these cavities had within them a void, in the rigorous sense of that word; rather, the cirri that can be observed across [the cavities] show that they are full of a less incandescent gas, but that sometimes can be very clearly seen turning in vortices and currents. Now if this is the case, what are these cavities if not simply spaces full of less luminous, and thus less incandescent, material? Signor Kirchoff prefers to imagine them as clouds or rather cooler masses. There is not, in fact, much to distinguish the two hypotheses, finally, provided that the terms are well defined. The difference is further diminished if we see the origin of such clouds that is attributed to vortices and cataclysms, which is the cause that we, too, have often attributed to the origin of the sunspots.

The only point of controversy that remains is to decide if that black [part] that is called the "nucleus" is a part of that general ground that remains beneath the photosphere, or if it was produced by the opacity of a cloud or a cooler mass which prevented the rays from the more luminous part beneath from reaching us.

This issue can only be resolved after scrutiny of the shapes and the phases of the sunspots themselves, and not in *a priori* fashion. Now the study of their shapes does not agree at all with that of clouds as far as we can judge from what happens in our atmosphere and what can reasonably be imagined to take place in an incandescent atmosphere such as that of the sun.

In fact, sunspots present themselves to us from the outset like black pores, in which it would not be difficult to recognize the idea of clouds, but soon enough all analogy vanishes. Because if the pore expands until it has the appearance of a spot, it can be observed that its edges are ragged, and the penumbra is formed *entirely of very fine rays converging towards the middle of the shape*. The nucleus does not always

present the outlines of the penumbra in rigorous fashion, as has been said several times, but rather, a *protruding* angle of the luminous material against the nucleus corresponds to an angle *sloping* into the penumbra, just as would a cascade of material that fell from the walls into the nucleus, which would leave a scarp (*talus*) whose slope would increase as greater amounts of material flowed. These are the phases of all sunspots as long as they are in the first stage, which seems to be that of formation and complete development, after which the phase of dissolution follows.

Thus it is apparent that this first phase cannot show us anything that is similar to what should happen when a cloud forms. The cloud should appear like a less luminous mass, and should be either decisively separated from other warmer ones as are our cumulus clouds, or shaded on the edges like our stratus clouds; that radiating shape and the appearance of currents running into a cavity and forming a distinct scarp will not ever be observed, in any guise, at least in what we can perceive and reasonably conjecture about our clouds. Whatever the theory of sunspots might be, their appearance must first of all be explained, and this appearance has yet to be explained by any theory that compares them to clouds.

When the sunspot has reached its full development, it shows vast black surfaces in which brilliant threads erupt like radiant torrents all around the photosphere, twisting in long contorted lines within the nuclei and breaking, as noted earlier. Now if we were to judge what is happening there on the basis of what happens in our atmosphere, these eruptions of warm masses within cold ones, occurring in such fashion that they remain distinct and constantly separated, cannot be observed by us at all, nor does it appear that they can be formed, because the cloudy opaque mass would either block them from our view, or the mass itself would diminish the light, thus cooling down the torrent that penetrates within [the nuclei] with that linear movement. Now as we have already observed several times, and as Signor Dawes has recently repeated as well (in the latest number of the *Monthly Notices*) the filaments of the photosphere that penetrate into the nuclei maintain an extremely brilliant light, as bright as the photosphere itself. Such a structure for the sunspots hardly confirms the idea of clouds.

When the sunspot is in the last phase of dissolution, the penumbra is less regularly radiated, and it seems formed of the thinnest and most tenuous part of the photosphere itself. In this phase it can be said that it has some analogy with clouds, but a theory, of course, must give an account of all the phases. There is, moreover, a circumstance of which the analogy with clouds explains nothing, and that is the presence of faculae that surround the sunspots.

These faculae are nothing other than the crests of the tempestuous waves excited by the photosphere, waves whose peaks emerge from the denser stratum of the solar atmosphere, as I have shown at length in other publications. They seem in fact formed by the photospheric matter that has been hurled about by the internal force that creates the sunspot. If the sunspot were nothing but a cloudy formation, there would be no explanation for why its contours should be agitated and violently thrown into disarray. Everything indicates that the sunspots are centers in which the temperature is less, and I have demonstrated as much with the thermoscope. But it is also clear that the source of these lacunae is rather an eruption coming from the inside of the nucleus, rather than a simple drop in temperature produced in the photosphere by factors analogous to those in terrestrial meteorology, which would be difficult to imagine in the sun, whereas internal eruptions cannot be avoided in a body placed in such conditions.

But there is something more: Herschel, in order to explain the penumbras proposed two layers to the photosphere, just as Signor Kirchoff proposed two layers of clouds which were always obliged to appear together, the one above the other. These two strata are surely a pure expedient to explain the penumbras, of whose composition we have already spoken, and which can be explained merely by proposing a simple photosphere with those features that are inseparable from fires of this sort. The hypothesis of the clouds has been frequently been raised, but always by those who either have not carried out much solar study, or who have undertaken it with imperfect and mediocre instruments. Thus this hypothesis has always been rejected by those who had at their disposal better means of observation. There is no need of the goal of proposing a less luminous nucleus, nor of that effort (as perhaps has been excessively emphasized) to revive the old fantasies of the habitability of the sun, because if the Creator had wanted to make this star habitable there would have been no need to place men of flesh and blood like us there, as they would be incinerated within a few seconds; nor is there any need to imagine, for that reason, that the black layer is like a tent to shelter such inhabitants from excessive rays. These matters might be useful to amuse the readers of a treatise of Fontenelle or of those who follow in his tracks. We are saying only that without contradicting the laws of physics, first, that the photospheric layer might possess a brilliance greater than that of the internal nucleus; second, that what we call "nucleus" absolutely does not need to be imagined either solid or liquid, but might even be gaseous alone, but more dense; third, that in spite of the proximity of the photospheric layer, it might have not only a different light, but also a different temperature; and fourth, that the appearances of the different shapes of the sunspots absolutely rule out cloud-like structures, and we see nothing in the sunspots that has sufficient analogies with the way in which our clouds are formed, or the changes through which they go.

We wanted, therefore, to say these things less to object to such a distinguished physicist, than to prevent science from taking a retrograde course, especially since history shows that persons of great authority in one branch [of knowledge] often drag along, under the weight of their opinion, those who are less experienced, even in matters where their studies are not

sufficiently deep and where they should not have such influence. We hardly pretend to have given a true theory of the sunspots, but we believe merely, as has been demonstrated, that the notion that they are clouds is surely one of the most infelicitous of hypotheses that can be imagined.

Acknowledgement

Chris Corbally and Sabino Maffeo of the Society of Jesus and the Vatican Observatory are recognized for their assistance in providing Professor Robitaille with copies of Father Secchi's original papers as they appeared in the *Bullettino Meteorologico dell'Osservatorio del Collegio Romano*.

Submitted on April 28, 2011 / Accepted on April 28, 2011
First published in online on May 07, 2011

On the Structure of the Solar Photosphere

A. Secchi

Observatory of the Roman College, Rome, Italy

Mary Posani (Department of French and Italian, The Ohio State University, Columbus, Ohio, 43221) and Eileen Reeves (Department of Comparative Literature, Princeton University, Princeton, New Jersey, 08544) provide a translation of Father Pietro Angelo Secchi's classic work "*Sulla Struttura della Fotosfera Solare*" as it appeared in *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 30 November 1864, v.3(11), 1–3. Secchi's paper was immediately translated into French by l'Abbé François Moigno appearing on December 22nd, 1864 (*Sur la structure de la photosphère du soleil. Les Mondes*, 1864, v.6, 703–707). Moigno's translation prompted significant interest in the nature of the Sun throughout Europe, with rapid claims of simultaneous discovery by Harvé Faye (Faye H. *Sur la constitution physique du soleil — premiere parti*. Les Mondes, v.7, 293–306) and others. In this article, Secchi reiterated that the photosphere was composed of solid corpuscles floating on the transparent atmosphere of the Sun. Secchi concluded that the body of the Sun was gaseous based on his visualization of solar granules or "*willow leaves*" described by Nasmyth (Nasmyth J. *On the Structure of the Luminous Envelope of the Sun — In a letter to Joseph Sidebotham. Memoirs of the Literary and Philosophical Society of Manchester*, 1862, 3rd Series, v.I, 407–411). Secchi also referred to Magnus' work on solid particles in the gaseous flame (Magnus G. *Notiz über die Beschaffenheit der Sonne. Poggendorff's Annalen der Physik und Chemie*, 1864, v.121, 510–512; also in French *Notice sur la constitution du soleil. Archives des science physique et naturelles (Genève)*, 1864, v.20, 171–175). The works by Sechhi, Nasmyth, Magnus, and Faye would dominate astrophysical thought for the next 50 years.

In the first number of the *Bulletin* of this year we mentioned in passing the discovery by Signor Nashmyth concerning the structure on the sun which he named *willow leaves*, and which was subsequently confirmed by other astronomers, who found it preferable to call them *grains of rice*, because of the greater irregularity in the forms of the oval parts that they observed.

We said then that we did not yet have the means to examine this structure, because we lacked an ocular which would enable us to examine the solar image without a black lens, or darkener, and with a full aperture of the objective lens, an apparently essential condition for the accurate observation of these details. Recently, however, having received such an ocular through the kindness of Warren de la Rue, we made, as soon as the atmospheric circumstances permitted it, a series of observations, which we now report, reserving for another occasion a more extensive exposition.

The eyepiece of which we speak was formed with a plate of dark red glass inclined at 45° that reflects a small portion of the luminous rays, while it transmits a large portion of the others, and especially those of caloric value; the axis of the eyepiece by consequence remains at a right angle with the axis of the telescope. In the more northern climates, and especially in England, one can with this simple means of weakening of the rays look at the sun without danger to one's eyesight; but for us it is impossible, and we must add a slight darkener. The method of observing the reflection was proposed also by Signor Porro and P. Cavalleri: they had used, instead, two glasses under the angle of polarization, where, because the reflecting planes, were at right angles, the light becomes tolerable to the eye without any other adjustment, and remains white.

With one or the other method, one can visualize the sun much better than with colored glasses; the light remains white, and thus we can distinguish many details that were lost in the older method. However the polarizing system, introducing a double reflection, requires a great perfection in the optical reflectors, and thus, it is difficult to apply it to large instruments, in which the reflective surfaces should be rather broad. Instead of the reflective colored glass, one can substitute a prism by reflecting rays on the hypotenuse externally (and not by total reflection). However one cannot use a simple strip with parallel faces, because the second surface would give an image that could disturb that of the first.

Applying therefore this new eyepiece to Merz's great refractor, maintaining an aperture of its nine inches, we could immediately recognize a structure that truly differed greatly from what is commonly observed. The bottom of the solar disc appeared to be formed of a fine black mesh whose links were very thin and full of bright points. It was not so much the shape of the grid that surprised us — for we had seen it also at other times with older methods — as its blackness, which was truly extraordinary. It was such that we suspected some illusion, but in concentrating on certain darker points and finding them of unchanging and precise forms, we no

longer remained in doubt about the reality of the aspect. Of this grid-like structure we can give an approximate idea in saying that the sun looked like a ordinary piece of rough paper seen through a strong microscope; on this paper the prominences are numerous and irregular, and where the light falls rather obliquely, the bottom of the grooves are almost black compared to the more elevated parts, which appear extremely white.

The grid-like solar structure seemed to us to offer nothing regular in those parts of the disc that are continuous, and thus the term *granular* appears very appropriate. Nevertheless, in the vicinity of the sunspots, that of *willow leaf* remains justified, because we actually see a multitude of small strips which terminate in rounded tips, and which encircles the edge of the penumbra and of the nucleus, resembling so many elongated leaves arranged all around. The granular structure is more visible near the spots, but it is not recognizable in the *faculae*; these present themselves like luminous clusters without distinguishable separation, emitting continual light without the interruption of dots or of that black mesh.

In the end, we have found the granular structure more notable and easy to distinguish in the middle of the disc than near the limb, and in the zones near the sun's equator, more than in the polar zones. The first [of these features] is without doubt an effect of the sun's refraction: in fact, the transparent atmosphere which encircles the sun must, because of its thickness and greater agitation, produce a greater confusion near the limb. We seemed to have recognized a trace of the effect of the refraction of this atmosphere in some of the systematic irregularities of the place of the spots near the edges, found by Signor Carrington in his admirable recent publication about sunspots. The polar regions, as is known, have a lower temperature and less agitation, and the spots do not appear there. This grid suggests that the spots are but an exaggeration of the minute holes which riddle the solar surface.

These are, in summary, the observations, which certainly raise a great number of questions. First of all, are these new findings? We believe that, in the end, these are the same granulations that have long since been pointed out by observers, under the name of "lucules" and "pores", and that with the new method they can better be distinguished. Because, since we can in this manner utilize a large aperture, the phenomenon of dilatation of luminous points or circles of diffraction that the objective lens forms are considerably diminished, and, as a consequence, we can better recognize the details, because each luminous center remains completely separated.

In the second place the rounded tips which surround the nuclei and the penumbrae are not new — at least not us — but rather are those that we have always indicated as evidence of the luminous currents that run to fill the emptiness of the spots. They are those types of currents that accumulate around the *nucleus*, and render the light appearing there greater in intensity than in the remoter regions of the penumbrae, just as the spokes of a wheel are more crowded together near the axle than towards the circumference. However, we must not omit the fact that with this means of observation, the appearance of a continuous current seems in many cases rather interrupted, and takes on instead the aspect of a multitude of torn fragments, or as Dawes says, of truncated straws that run towards the nucleus. But in any case, the more we study with attention these phenomena, the more it is *unacceptable* to us the idea that the spots are clouds. We do not hesitate to say there is still much that is mysterious in this structure, but certainly it has nothing to do with clouds, unless we wanted to say that the clouds are rather what form the luminous element, and that this incandescent material rushes in *cumulus* and in *cirrus* in the void of the spots, as we see sometimes in our atmosphere the *cumulus* and the *cirrus* run and fill in voids [in the sky].

Indeed this appearance suggests to us what is perhaps a bold hypothesis. As in our atmosphere, when it is cooled to a certain point, there exists a fine substance capable of transforming itself in fine powder and of forming clouds in suspension, (water transforming into so-called "vesicular" vapor or into small solid icicles), so in the enflamed solar atmosphere there might be an abundance of matter capable of being transformed to a similar state at the highest temperatures. These corpuscles, in immense supply, would form an almost continuous layer of real clouds, suspended in the transparent atmosphere which envelopes the sun, and being comparable to solid bodies suspended in a gas, they might have a greater radiant force of calorific and luminous rays than the gas in which they are suspended. We may thus explain why the spots (that are places where these clouds are torn) show *less* light and less heat, even if the temperature is the *same*. The excellent results obtained by Magnus, who has proved that a solid immersed in an incandescent gas becomes *more radiant* in heat and light than the same gas, seem to lend support to this hypothesis, which reconciles the rest of the known solar phenomena.

In the third place, one will ask oneself if such appearances are constant. It seems that we should say yes, since it has been discussed for two years, but to observe them takes no small practice and good instruments. Ours was extremely well made, with a red lens from England, but it showed little resistance to our [Italian] sun, and exploded into many pieces. Now we have substituted a prism, but it emits too much reflected light, and its surface is perhaps not perfectly polished. Nonetheless, we continue to see with clarity a grid and the other phenomena mentioned above. But the principal obstacle is the agitation of the air, which by mixing all these small shapes, makes a general confusion and flattens everything, for which reason they are only seen intermittently on those days that are anything short of perfectly calm. However, by moving the telescope slowly we can see the granulations much more easily than when we hold it fixed, and once they are recognized, it is easy to follow them and to study their forms.

May these indications suffice for now; the numerous other questions raised by this new method of observation and by this structure will be resolved with time. For now it is certain that this mode of observation can be said to have truly been a new conquest of practical astronomy.

Acknowledgement

Chris Corbally and Sabino Maffeo of the Society of Jesus and the Vatican Observatory are recognized for their assistance in providing Professor Robitaille with copies of Father Secchi's original papers as they appeared in the *Bullettino Meteorologico dell'Osservatorio del Collegio Romano*.

Submitted on April 28, 2011 / Accepted on April 28, 2011
First published in online on May 07, 2011

A Note on the Constitution of the Sun

H. G. Magnus
University of Berlin, Germany

Patrice Robitaille (TAV College, Montreal, Canada) provides a translation of Heinrich Gustav Magnus' classic work *Notiz über die Beschaffenheit der Sonne*, as it appeared in March 1864 within *Poggendorff's Annalen der Physik und Chemie*,1864, v.131, 510–512. The article had previously been translated into French: *Notice sur la constitution du soleil. Archives des science physique et naturelles (Genève)*, 1864, v.20, 171–175. This work formed the basis of the present translation. Heinrich Gustav Magnus (May 2, 1802 – April 4, 1870) was a professor at the University of Berlin and had studied in Paris under Joseph Louis Gay-Lussac. He would count amongst his students Wilhelm Beetz, Hermann Helmholtz, Gustav Wiedemann, John Tyndall, Rudolph Weber, and Adolph Wüllner (Heinrich Gustav Magnus, *Platinum Metals Review*, 1976, v.20(1), 21–24). In his *Notiz*, Magnus demonstrated that the addition of sodium hydroxide to the gaseous flame resulted in a tremendous increase in luminosity. Magnus' work would inspire Father Secchi to propose that the Sun was a gaseous globe whose photosphere contained condensed particulate matter (Secchi A. *Sulla Struttura della Fotosfera Solare. Bulletino Meteorologico dell' Osservatorio del Collegio Romano*, 1864, v.3(11), 1–3; English translation in *Progr. Phys.*, 2011, v.3, 30–32). Magnus' report on the constitution of the Sun would continue to impact solar physics for two generations.

Already in 1795, W. Herschel* advanced the idea that the Sun is formed of an obscure nucleus surrounded by an atmosphere or photosphere from which light and heat are emitted. Between this photosphere and the nucleus, he also admits the presence of a reflective atmosphere whose reflection prevents the light of the photosphere from reaching the nucleus. Arago† in exposing this hypothesis which he gives as generally accepted‡, remarks that the photosphere determines the outer edge of the Sun, but that the photosphere is itself surrounded by a diaphanous atmosphere; he comes to this conclusion through the observation of the protuberances [flares and prominences] during total eclipses of the Sun. Herschel§ says that the photosphere is neither liquid nor gaseous, but that it is made up of luminous clouds. According to our current knowledge of the radiation of light and heat, it is difficult to admit that the photosphere, from which solar heat emanates, does not heat to the point of incandescence the nucleus that it surrounds. The intermediate reflecting atmosphere, whose existence was assumed, could very well stop the passage of light but not the progressive heating of the nucleus. It is therefore with reason that Mr. Kirchhoff¶ says that this hypothesis which was devised to explain sunspots, is in such total contradiction with our knowledge of physics, that we should reject it even if we cannot come to make comprehensible, in another way, the phenomenon of sunspots.

Mr. Kirchhoff was guided by his research on the solar spectrum to admit that the Sun consists of a solid or liquid nucleus, brought to the highest incandescence and surrounded by a diaphanous atmosphere with a slightly inferior temperature.

I do not know that we have as yet deduced from the nature of the heat that emanates from the Sun, a conclusion on its constitution; we could but mention the observations of Reverend Father Sechi‖ relating that the poles emit less heat then the Sun's equator. Some of the experiments that I have conducted on calorific radiation, allow, I think, for new views on the constitution of this celestial body.

If we observe the heat that emanates from a non-luminous gas flame, and if we introduce a bit of sodium hydroxide which, as we know, renders it extremely luminous, we see at the same time, that the calorific radiation increases. The experiment was carried out in such a way that we were always comparing a predetermined place of the sodium hydroxide flame, with the same place of the non-luminous flame, and this in such a way that the sodium hydroxide introduced into the flame, could not radiate over the thermo-electro battery used for observation. Evidently, in this case, part of the heat of the flame was used to bring to incandescence or to vaporize the sodium hydroxide and the platinum blade on which it was found in such a way that, in the end, the flame had a lower temperature than before when it was not luminous, and yet it emitted about a third more of the heat that it had previously.

It can be that the sodium hydroxide was contained within the flame in a state of vapour or that particles removed from that body that augmented the illuminating power. Whatever the case may be, I choose, to shorten the discourse, the des-

Philosophical Transactions for 1795, page 42.
†*Astronomie populaire*, Vol. II, page 94.
‡*Ibid.*, page 143.
§*Philosophical Transactions* for 1795, page 71.
¶*Denkschriften der Berliner Acad. Der Wiss.*, 1861, page 85.
‖*Comptes rendus de l'Acad. Des sciences*, Vol. XXXV, page 606 and Vol. XXXVI, page 659.

ignation of sodium hydroxide vapour.

By introducing in the place of this vapour a platinum disk in the same area of the flame that was being studied, the heat that the flame emitted became even more considerable than previously recorded. The platinum blade evidently removed from the flame even more heat than the sodium hydroxide, but it, however, radiated even more. With the blade that I was using and whose diameter was 55 mm, the radiation became nearly twice as strong then when the flame did not throw off any light. We did not observe any fundamental difference by making the blade thicker or thinner, so long as the diameter remained the same.

But if, instead of making the blade thicker, we covered it with sodium carbonate, then the radiation increased again considerably; it became fifty percent stronger than with the platinum blade without any sodium hydroxide.

The radiation would rise even more, when, apart from the platinum blade being covered with sodium hydroxide, there was also sodium hydroxide vapour in the flame, this being obtained by introducing in the lower part of the flame some sodium hydroxide on a platinum blade, in the same way as was done previously, that is to say without having the sodium hydroxide radiating over the battery.

In this case, the flame being completely filled with sodium hydroxide vapour coupled with the platinum blade covered with sodium hydroxide, the flame emitted close to three times more heat than the flame that was not luminous. Lithium hydroxide and strontium hydroxide behaved like sodium hydroxide.

These experiments demonstrate that gaseous bodies emit far less heat than solid or liquid bodies; and that, by consequence; one cannot assume that the seat of solar heat resides in a photosphere made up of gas or vapours. They also demonstrate that, and this is especially striking, that incandescent sodium hydroxide has a much greater radiative power for heat than platinum at the same temperature.

Also, they demonstrate that sodium hydroxide vapour or sodium hydroxide particles absorb only a small part of the heat emitted by incandescent solid or liquid bodies. In fact, the radiation of the solid body in the flame filled with sodium hydroxide vapour was, it is true, always smaller than the sum of the radiations of the solid body alone and of the vapour introduced alone in the non luminous flame, but the difference was small.

This manner in which incandescent liquid or vaporous sodium hydroxide behaves confirms in a striking way the views of Mr. Kirchhoff on the nature of the Sun.

Submitted on May 13, 2011 / Accepted on May 18, 2011
First published online on May 31, 2011

On the Physical Constitution of the Sun — Part I

H. Faye

Académie de Nancy, France

Patrice Robitaille (TAV College, Montreal, Canada) provides a translation of Hervé August Etienne Albans Faye's classic report *Sur la constitution physique du soleil*, as it appeared in February 1865 within *Les Mondes*,1865, v.7, 293–306. Hervé Faye (October 1, 1814 – July 4, 1902) led a distinguished life, both in science and public service. He was widely regarded as one of the premier astronomers of his day. He had studied under the great François Arago. In 1843, he became a *Chevalier de la Legion d'Honneur* and, in 1877, served as the French Minister of Education (Catholic Encyclopedia, 1913). Faye's report *On the Physical Constitution of the Sun* was a crucial milestone in the history of astronomy. It was through this paper, that the Sun became viewed as devoid of a distinct surface. The work was also interesting as it presented Faye's early conception of the gaseous Sun. In addition, through its submission, Faye had sought the approbation of Father Secchi relative to claims of simultaneous discovery (see P. M. Robitaille. *A Thermodynamic History of the Solar Constitution — I: The Journey to a Gaseous Sun. Progr. Phys.*, 2011, v.3, 3–25). Faye's work would continue to impact solar physics until the 1920s.

Why do astronomers have so much trouble describing the physical constitution of the Sun? Why so many contradictory conjectures? One tells us that the Sun is an opaque globe, obscure, cold like ours, perhaps even inhabited, but surrounded by a radiant aureole, from which is emitted the heat and the light which, for thousands of centuries, has given life and movement on our little world of planets. Yet another affirms that it is a liquid globe, incandescent, surrounded by a vast atmosphere where float clouds of iron, sodium and magnesium vapor, etc.

It is in such a way that the sciences make their first appearance when they possess but a small number of facts and laws. The human spirit needs conjectures in order to take interest in the things that are beyond reach. But the question of the Sun cannot remain where it is after two and a half centuries of diligent observation. We have gathered, on this matter, the main elements of a rational solution; it is now time to address it.

What is the difference between a conjectural solution and a rational solution?

The first is quite simple; you have observed two or three facts: to explain them, imagine as many particular entities as there are facts, and try to coordinate them in a way to avoid that they contradict each other. Before the telescope, the only thing we knew about the Sun was its extremely powerful heat and its unwavering brightness; the conjecture consisted to say that this celestial body was formed of a subtle element, incorruptible, infinitely more noble than our terrestrial flames which smoke and die out miserably. Also, the discovery of sunspots would strongly appall the partisans who believed that the heavens were incorruptible; and when Father Scheiner, to whom we owe such remarkable work on these phenomena, went on to mention these to his superior, the latter replied to him: "I have read and reread Aristotle, but I haven't found anything there touching the things you tell me; go, my son, hold your spirit to rest; there are no spots on the Sun other than those that are created by the defects of your eyes or of your telescopes".

But the conjecture had to yield before the facts. These facts, are described here in all their simplicity: black spots are produced on that shining pool of fire; they are born, take about two weeks to cover the distance of the solar disk, and then pass over to the other side; we see them again at the end of another two weeks; sometimes they persist for months; normally, they disappear after a few weeks. These spots really look like holes; we can even distinguish, using powerful telescopes, a less brilliant part that typically resembles the embankments of these holes. The bottom seems completely black. Black holes in a pool of fire! It is apparent that the brilliant part is only a rather thin envelop of a very mobile luminous fluid, covering a black core, and here lies the second conjecture. We have lived a long time on that one and it has its merit. The preceding one itself, I mean the incorruptibility of celestial bodies, also had its own merit, since it represented a great fact, still true today, as in the time of the scholastics.

Lastly, in more recent times, a capital discovery revels the minds with much admiration: The rays of the solar spectrum are explained; we reproduce them in the laboratory by placing metallic vapors on the path of a beam of light that emanates from an incandescent solid or liquid.

Let us conjecturally transport this experience to the sky: the Sun will become an incandescent solid or liquid surrounded by a vast atmosphere of metallic vapors. But what about sunspots? How can black holes form themselves in a liquid or solid? Here, we must avoid an absurdity; the spots will be produced by something exterior, precisely by the clouds of that atmosphere, clouds formed of metallic vapors that begin

to condense. Whatever can be said on the matter, this latest idea, which seems to violate all facts, except one, nevertheless answers to one of the most admirable discoveries of our time, that of spectral analysis, which permits the pronouncement, by the appearance of a light, on the chemical nature of the environment through which it has travelled.

During this time, the facts were multiplying, I am not saying at the time of Aristotle when we did not have telescopes, but since Fabricius, Galileo, and Father Scheiner. Today, the enumeration of observed facts offers a magnificent total. We must ask ourselves, I repeat, whether, in the presence of these facts, it is not time to renounce conjectures and to try a little simple reasoning. This second method is that which definitively constitutes science: it only comes after the first, but it must also have its turn.

Here, it is no longer a question of guessing, but of linking the phenomena through laws known in the physical world to some simple and very general fact that we would not be tempted to set aside. I do not know if I have succeeded, I am certain, at least, that the time has come and, since it is a question of pure logic, another, reasoning better, will succeed if I have failed.

My starting point will naturally be solar heat. Everything proves to us that this heat must be enormous; it must enormously surpass the highest temperatures that we can produce in our laboratories. However, the former suffices already to break down a large number of bodies. We must therefore consider chemical phenomena as being capable of occurring beginning at a certain temperature scarcely remote from those we can produce, but not above them. Above them, the elements mix, but do not combine. In the same manner, the phenomena of electricity, magnetism, life, occur at a certain temperature, but not above it. There is reason to believe that the Sun is at a temperature of universal chemical and physical dissociation, that its heat much surpasses all affinities, all molecular attractions, in such a way that its entire mass reduces itself to a gaseous mixture, to a true chaos of entirely separated atoms. That is my starting point, of which the complete justification, based on the dynamic theory of heat would require much too lengthy developments. I then place on one side the most characteristic known facts, on the other the consequences of my premises; if the starting point is accepted, if the facts can be successively identified with the consequences, we will have drafted a theory and no longer a conjecture.

This mass is undergoing cooling, since nothing comes from the outside to restitute the heat that it throws off daily into space, the stellar radiations being extremely weak; from there the successive phases which are convenient to analyze first.

In fact, the enormous heat that we have just mentioned is that of the entire mass; at the surface, there where cooling operates with the most energy, it can fall far below the internal heat, and make way for the initiation of chemical activity. Is this deduction true, can it be applied to the Sun? To find out, let us consult the facts. The heat emitted has been measured: it has been calculated that it does not exceed 30 or 40 times the heat contained in the furnace of a locomotive when it actively draws energy. On the other hand, the most intense heat furnaces produced by man do not emit a light incomparably weaker than solar light. We can therefore admit that, on the surface, chemical actions start to produce themselves, at least those that give birth to the most stable components. There are two ways, in fact, to have affinities react in a mixture of gas and vapors; by heating, if the mixture is cold; by cooling, if the mixture has gone beyond the temperature of dissociation.

Thus, in this environment, particulate clouds will be produced that will no longer be gaseous, but liquids or solids, like magnesia in a mixture of vaporous oxygen and magnesium and, in another sense, like the carbon in our lighting flames. Now these particles, becoming incandescent, will radiate enormously more than the gaseous environment itself, at the same temperature, because their emissive power is much superior to that of elementary gases or vapors. As a result, by the sole fact of superficial cooling, any gaseous mass primitively brought to a temperature of dissociation will surround itself at the surface with a continuous or discontinuous luminous cloud.

To these conclusions answers, item by item, as we shall see, the photosphere of the Sun.

There is, however, one difficulty. In a hot gaseous mass, isolated in space and which is cooling, there can and there must be established after a certain time, and following interior movements, a certain equilibrium that temporarily opposes the transport of some portion of the mass from one layer to another. Admitting, therefore, that chemical action occurs at a given instant in the exterior layers, following this cooling, how would it be maintained? How could the photosphere, which is produced momentarily, renew itself continuously and regularly? Here is the answer. The non gaseous particles that form the photosphere's luminous clouds are much heavier than the gaseous environment from which they are born; they will obey the attraction of the entire mass, and will fall vertically until they reach a layer that is hot enough to reproduce the dissociation of their elements. But then; in that layer, the gases and vapors due to this dissociation will break the equilibrium and will force a certain part of the mass of this layer to elevate itself to superior layers. From this, results a double incessant current that would produce itself only on long intervals and in a tumultuous manner, if the mass remained gaseous everywhere, if the chemical activities did not intervene to modify all at once the density of the superficial parts. This double current therefore incessantly brings to the surface part of the internal heat that is dispensed rapidly, thanks to chemical activity; while the incandescent particles, because of their excess density, fall once again within the deeper layers and lower, little by little, the temperature. There lies, to my liking, the rational explanation of that marvelous

constancy of solar radiation, first phenomenon that hit the ancient [philosophers], whose long-lasting conjectures have never tried to take into account. How could the Sun, considering only historical times, support its enormous radiation with such a luminous envelop, thick of only a few leagues, being the seat of the most curious phenomena? The combustion of all the elements composing the Sun would not represent heat capable of supplying this radiation during half of that short period. Do you adopt the second conjecture, that of Mr. Kirchhoff? The thing would become even less possible still, because a liquid envelope would be quickly cooled; it would encrust itself at the surface, while the interior would maintain a high temperature that would have no other outlet but the weak conductibility of the outer crust. Conversely, the rational explanation of the photosphere gives for the energetic constancy of the radiation the only admissible reason, by showing that the entire mass participates in this heat expenditure and not only the superficial area. It must be remembered that the entire mass is enormous and that the originating temperature is equally enormous.

If I insist on this point, it is because here lays the heart of the problem. Everything else will easily follow if, on this point, one is willing to permit me to advance my cause. This old problem that the ancient school had resolved in its own way by proclaiming the incorruptibility of the heavens, was simply set aside by modern thinkers, until the creators of the dynamic theory of heat decided to revive the discussion. But their solution, so scholarly and so ingenious, was just one more conjecture: they believed they had to invent an artificial means to maintain this enormous caloric expenditure that equates to the incessant production of a 75,000 horse-power force for every square meter of solar area, while it suffices to represent a mode of cooling such that the internal mass is constantly called to supply to the superficial area the heat that it emits.

So then the exterior surface of the Sun, which from far appears so perfectly spherical, is no longer a layered surface in the mathematical sense of the word. The surfaces, rigorously made up of layers, correspond to a state of equilibrium that does not exist in the Sun, since the ascending and descending currents reign there perpetually from the interior to the superficial area; but since these currents only act in the vertical direction, the equilibrium is also not troubled in that sense, that is to say, perpendicularly to the leveled layers that would form if the currents came to cease. If, therefore, the mass was not animated by a movement of rotation, (for now we will make of it an abstraction), there would not be at its heart any lateral movement, no transfer of matter in the perpendicular direction of the rays. The exterior surface of the photosphere being the limit that will attain the ascending currents which carry the phenomenon of incandescence in the superior layers, a very-admissible symmetry suffices in a globe where the most complete homogeneity must have freely established itself, to give to this limit surface the shape of a sphere, but a sphere that is incredibly uneven.

This limit is in any case only apparent: the general milieu where the photosphere is incessantly forming surpasses without doubt, more or less, the highest crests or summits of the incandescent clouds, but we do not know the effective limit; the only thing that one is permitted to affirm, is that these invisible layers, to which the name atmosphere does not seem to me applicable, would not be able to attain a height of $3'$, the excess of the perihelion distance of the great comet of 1843 on the radius of the photosphere.

If you compare now these deductions to the best known facts of detail, you will find a remarkable agreement. The incessant agitation of the photosphere, the black points or rather the little interlaced black lines that cover the surface, the spots and the faculae are easily understood if we refer to the action of the vertical currents that we have just described. What shines in fact are the products of the chemical activity, that occurs in the photosphere on matter that is constantly renewed by the currents, and not the gaseous environment where these incandescent phenomena take place. To properly understand this difference, it would suffice to observe, through one of those obscuring glass plates that astronomers use to observe the Sun, the flame of pure hydrogen, or the one produced by a Bunsen burner, next to a flame produced by magnesium vapor. The first would be completely invisible, that is to say black; the other would be as white as snow. If, therefore, for one reason or another the incandescent clouds of the photosphere come together in a given place, there the visual ray will only meet but the general gaseous mass of the Sun endowed with a very weak emissive power, while a little further the photosphere will appear with its intense radiation and dazzling brightness. Father Secchi, recently came to a similar explanation of sunspots which makes me hope that the ideas I have just presented on the formation of the photosphere will meet his approbation. As for the faculae, there is nothing simpler assuredly that such level differences at the extreme limit of our ascending currents, and nothing so difficult to understand for those who admit the liquid photosphere. Persistent ridges of 100 or 200 leagues high on the extreme surface level of a liquid layer are not easy to justify.

But the high point of this theory, is the reconciliation of the two famous and contradictory experiments of Arago and Kirchhoff. Basing himself on the polariscopic analysis of the light of the Sun, Arago concluded that the photosphere had to be gaseous; basing himself on spectral analysis, Mr. Kirchhoff concluded that the photosphere is solid or liquid. The only way to have these opposed conclusions agree is to admit the photosphere I have proposed. Non gaseous but incandescent particles, floating like a cloud in the midst of a gaseous environment, would in fact emit natural light under all angles of incidence; they would also emit rays of all refrangibility with the exception of those that the gaseous environment interposed between the particles is capable of absorbing. The second point is the only one that needs a few developments:

the light so emitted is not purely superficial, it comes from a great depth; by consequence the largest part of the rays incurs on the part of the general environment, a very strong absorption. It would be different if the light was emitted only from the superficial area, then an exterior environment would be required, interposed between that superficial area and us, in order to produce the required absorption; as it is seen in the vast atmosphere that Mr. Kirchhoff places around the Sun; but then the spectrum of the outer edges of the Sun would be considerably different from the spectrum of the center, because of the thickness of the atmosphere that would be much greater on the edges than in the center. However, the experiment by Mr. Forbes and the more recent and even more decisive work of Mr. Janssen establish that there is no difference between the two spectra; so the absorption comes mainly from the cause that I have assigned, and far less of the layers exterior to the photosphere, these being in reality but the far restrained continuation, in my opinion, of the general gaseous mass. It suffices to admit that the effective depth is the same in all directions where emission operates, and that it is then the same in the center and on the edges of the visible disk, a result to which I concluded some years ago through many other considerations.

To this gaseous mass, let us restitute now the more or less slow rotational movement it must have acquired, at the same time as its heat, through the gathering of the matter that constitutes it; the ascending and descending currents will incur, because of this rotation, a certain deviation. Originating from a great depth, the ascending currents reach the surface with a linear speed which is reduced since the rays of their primitive parallels were smaller. The photosphere whose matter is constantly renewed by these currents, must therefore be behind on the general movement of rotation; on the other hand, the theorem of areas requires that the sum of the projections of the areas described at a given time by the vector rays of the molecules remain constant, no matter the interior movements. This means that if the exterior layer is lagging the general angular movement, there will be, through compensation, an advancement of this angular movement for a few interior layers, and this is immediately understood, because the ascending currents cannot exist without, at the same time, the existence of descending currents that carry back the superficial materials towards the interior with the excess linear speed due to their greater parallels. Falling towards the center, this matter will therefore transfer this excess of speed to the layer where it has just incurred the dissociation of its elements. From this, there will be two layers to distinguish: a superficial layer that lags behind, and an internal layer that runs ahead of the angular movement that the entire mass would take if vertical equilibrium came into being. But some zone, in a rotating fluid must tend to approach the axis if it is lagging behind, and distance itself from it if it is running ahead on the speed of the general movement; so that the exterior layer will have a tendency to flow little by little toward the poles, while the interior layer which is in advance, will express the opposite tendency and elevate itself toward the equator. From this results a significantly complex modification of vertical currents that we first considered in all their simplicity, and I imagine that things will occur as if the interior layers from which they emanate were a lot closer to the center toward the poles than at the equator itself. If this deduction were founded, and one cannot argue with the fact that the term *layer* has a variety of meanings, it would manifestly result that the superficial rotation should vary from the equator to the poles and slow down, more and more, without, however, that the exterior feature would substantially cease to differ from the primitive spherical form.

Thus, the photosphere would be constituted of successive zones, parallel at the equator, animated by a decreasing angular speed in a way that is more or less continuous from the equator to the poles, while the inverse would produce itself in a certain deeper layer. In this complex phenomenon, that would be impossible to subject to calculation, the movements would operate mainly in the direction of the parallels either to the opposite, either in the direction of the general rotation, without this bringing about strongly marked currents in the direction from the equator to the poles or inversely. This is, therefore, a considerable phenomenon, a very special mode of troubled rotation that the planets could not present an exact equivalent, since the conditions there are so different.

In the case of the planets, in fact, one must make a distinction that does not need to be made in the case of the Sun, between the solid body of the planet and its atmosphere: the solid body turns altogether; it would be the same for the atmosphere, if an exterior action, the solar heat, did not intervene at every instant. Equilibrium therefore cannot exist in that atmosphere, but the phenomena that are produced there being regulated mainly by a notable difference in temperature between the poles and the equator, the movements being hindered by the presence of an unchanging solid or liquid surface (the surface of the solid globe on which rests the atmosphere), it is principally produced a lateral call of the atmospheric masses in the direction of the meridians, from the poles toward the equator. A superior counter-current is established in the same time in the inverse direction, in the layers that are further from the ground. Nothing like this happens on the Sun because the presence of the photosphere does not interrupt the continuity of the [central] mass, because there is no resistant ground to deviate the currents, because there is no exterior cause to trouble the equilibrium of the layers in the lateral direction. In order to illustrate the difference, I would say that, in the photosphere, the rotation only generates currents that are approximately directed along the parallels in the inverse sense of the rotation, while that, on the planets, the currents in the inverse sense of the rotation result as a medial or indirect effect of the superficial transfer of air masses in the direction of the meridians.

In short, it results, because of the appearance and the up-

holding of an atmosphere, in a gaseous mass animated by a rotational movement, that the surface must be delayed relative to the internal mass, in such a way that the superficial currents act only in the direction of the parallels save a slightly marked tendency toward the poles; and that this superficial delay must go increasing from the equator to the poles following a certain law that would be impossible to assign ahead of time, but of which we know this, that the direction of the rotational axis must not be substantially altered. Let us examine if the facts are in agreement with these consequences.

Here, it is good to restate things from a higher perspective. The astronomers naturally started by treating the Sun's rotation with the simplest hypothesis, that is to say, admitting that the entire mass turns as a single unit altogether, as if it consisted of the Earth or any other planet. In that case, the accidents of the surface would be animated with the same angular speed, no matter what was their position next to the pole or next to the equator, above the visible surface or below it. But this conjecture, the basis of all the work carried out in that sense from 1610 to 1840, was too far away from the truth for us to approach satisfactory results. If the astronomers generally agreed on the direction of the axis of rotation, they would reach the most discordant results concerning its duration. In the end, Delambre, discouraged by this failure, would console himself by saying that, after all, the subject had little importance, that it was good for training beginners. That was disregarding too hastily one of the most important phenomena of our solar world and one of the most verified laws in the history of the sciences, that is to say that all well-observed discordance carries with it the seed of a discovery. Finally, an astronomer was able to rid himself of these preconceived ideas in order to observe the phenomena for and in themselves. Mr. Laugier observed that every spot gave, so to speak, a specific value for the duration of the rotation: for 29 spots observed by him with all the refinements of precision, he observed that the completed rotations varied from 24 to 26 days, a difference far superior to the little uncertainty of the observations. This could mean two things: either the spots were animated by strong proper movements, or the successive zones of the photosphere did not possess the same rotational movement. Mr. Laugier left these things in that state, but he broke the ice, as we commonly say it, without mentioning the definitive elements that he had given to science for the direction of the solar axis. What needed to be done in order to pursue the work so nicely initiated? The spots had to be observed in a continuous manner, someone had to devote himself exclusively to this work for many long years, in order to discover the law of these specific variations; above all, a less dangerous method of observation for the eyes had to be devised by sacrificing partly the precision of the measurements.

That is what undertook Mr. Carrington, already known by astronomers through the great breath and extreme value of his work. Seven years and a half of continuous observation, 5 290 solar spot positions with the enormous quantity of drawings needed to conduct the discussion; there is the material that he accomplished. The definitive result can be formulated in the following manner: the determined rotation by the movement of sunspots is the same for all of the spots located at the same latitude, be it at the north, be it at the south of the equator, but it varies in a continuous fashion with latitude and becomes slower and slower towards the poles. Mr. Carrington tried to represent the complex phenomenon empirically with the following formulation: The duration of the rotation, obtained by dividing 216 000 by the movement of a spot expressed in minutes, this diurnal movement is equal to $865' - 265' \sin \frac{7}{4} l$, l designating the heliocentric latitude of the spot, and the quotient representing the average solar days. I do not know of any modern discovery that treats a matter more considerable than this one. We will not suppose, in fact, that the spots, simple clearing in the photosphere, could have such rapid proper movements (2 000 leagues per day at the 35th degree, for instance) and that they displace themselves this way within the environment where they are formed. A clearing, in a cloudy sky, can certainly displace itself and can displace itself at a great speed, but with the condition of being carried by the general movement of the ambient mass, which does not exclude specific modifications in the form and in the situation. We could not refuse ourselves to conclude from the nice work of Mr. Carrington that the photosphere moves with a varied angular movement whose slowness increases from the equator and up to the 15th degree and beyond and that this movement constitutes a mode of rotation quite different from that of the planets and their satellites.

Can this movement be assimilated to the trade winds and to the monsoons of our atmosphere? Observation answers negatively [to this question]. Trade winds originate from the transport of polar air masses toward the equator; the masses animated by a speed of rotation that is linearly less than the parallels met successively, appear to be blowing in the inverse direction of the terrestrial rotation, but here the essence of the phenomenon is not in the east-west sense of our trade winds, but the north-south direction (for our hemisphere); the first is but a consequence of the second, and the east-west movement would not exist if the movement from the north to the south disappeared or became too weak. However, on the Sun, we do not find any constant trace of this general movement from the poles to the equator, but rather an inverse tendency, starting from the 15th degree of latitude, from the equator to the poles, the identical tendency to the one that results from our above reasoning. Hence, the analogy that was naturally suspected at first does not exist, and we essentially remain before a new perturbation in a movement of rotation. It is up to the reader to decide if this great and beautiful phenomenon corresponds to the consequences that we have deduced from our theory.

One will surely note that these consequences end up being a little uncertain; this occurs because the facts themselves are not completely known. The formula provided by Mr. Car-

rington is purely empirical; the spots are so rare in the first degrees of the equatorial zone and from the 55th to the 50th degree of latitude, that the relative determinations in these zones are far from deserving the degree of confidence that can be given to the rotations concluded for the zones found between 5° and 35°. There is therefore a new work to undertake to complete the work of the English astronomer, but I do not think we can fully succeed without the help of photography whose introduction in the observatories is now a matter of factual use with our neighbors across the English Channel.

In short, conjectures no longer serve progress; they can only hinder it from now on. To the very simple idea associated with the cooling of a gaseous mass brought to a temperature such that its elements find themselves in a state of complete dissociation, except at the surface, where the chemical forces begin to exert themselves it is possible to logically link:

The constancy and the long duration of solar radiation;
The production and the maintenance of the photosphere;
The apparent contradictory experiments of Arago and Kirchhoff;
The explanation of sunspots and faculae;
And the mode of rotation particular to the Sun.

P.S. "I ask for permission to indicate here a coincidence or rather a remarkable agreement between the diverse conditions of organic life on the surface of the planets and our solar world. These conditions are of two kinds: 1) the mechanical stability of the system; and 2) the permanence of solar radiation. Either one or the other stability, even though they are of very different types, essentially rest on the enormity of the mass of the central celestial body".

Acknowledgement

Professor Robitaille would like to thank Laura Lingham and Judy Willis from the John A. Prior Health Sciences Library at The Ohio State University for their assistance in locating a complete version of Faye's *Les Mondes* article.

Submitted on May 13, 2011 / Accepted on May 18, 2011
First published online on May 31, 2011

A Thermodynamic History of the Solar Constitution — II: The Theory of a Gaseous Sun and Jeans' Failed Liquid Alternative

Pierre-Marie Robitaille

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

In this work, the development of solar theory is followed from the concept that the Sun was an ethereal nuclear body with a partially condensed photosphere to the creation of a fully gaseous object. An overview will be presented of the liquid Sun. A powerful lineage has brought us the gaseous Sun and two of its main authors were the direct scientific descendants of Gustav Robert Kirchhoff: Franz Arthur Friedrich Schuster and Arthur Stanley Eddington. It will be discovered that the seminal ideas of Father Secchi and Hervé Faye were not abandoned by astronomy until the beginning of 20th century. The central role of carbon in early solar physics will also be highlighted by revisiting George Johnstone Stoney. The evolution of the gaseous models will be outlined, along with the contributions of Johann Karl Friedrich Zöllner, James Clerk Maxwell, Jonathan Homer Lane, August Ritter, William Thomson, William Huggins, William Edward Wilson, George Francis FitzGerald, Jacob Robert Emden, Frank Washington Very, Karl Schwarzschild, and Edward Arthur Milne. Finally, with the aid of Edward Arthur Milne, the work of James Hopwood Jeans, the last modern advocate of a liquid Sun, will be rediscovered. Jeans was a staunch advocate of the condensed phase, but deprived of a proper building block, he would eventually abandon his non-gaseous stars. For his part, Subrahmanyan Chandrasekhar would spend nine years of his life studying homogeneous liquid masses. These were precisely the kind of objects which Jeans had considered for his liquid stars.

1 The search for a continuous thermal spectrum: Carbon particles on the Sun?

Consider particulate matter floating on a gaseous globe. Such was the idea advanced by Father Angelo Secchi and Hervé Faye as they described the photosphere of the Sun [1]. But what was this particulate matter? For Faye, a subtle allusion was made to carbon within the gaseous flame [2, p. 296]. As a result, the marriage between Faye's model and graphite was almost immediate. Graphite, or at least some form of condensed carbon, remained on the surface of the Sun until the 1920's. Even the pioneering treatment of a gaseous Sun, by Jonathan Homer Lane, referred to the carbon envelope of the photosphere, as demonstrated in Section 2.2. Thus, it was only through Eddington and his inception of a *fully* gaseous Sun [3] that particulate matter was finally removed from the photosphere.

If carbon played a pre-eminent role in solar theory, it was because of the need to understand the continuous spectrum of the photosphere. On Earth, only graphite and soot were known to produce such a spectrum. As the common form of condensed carbon, graphite possessed outstanding refractory properties. The material did not melt. Rather, it sublimed at extreme temperatures [4]. It seemed to be the perfect candidate for introducing condensed matter on the Sun in order to generate the solar spectrum. Moreover, from the earliest studies on thermal radiation [5, 6], graphite and soot played a dominant role [7]. Balfour Stewart [8] who, along with Gustav Kirchhoff [9], was one of the fathers of thermal emission, emphasized the crucial role of carbon in heat radiation: "*Indeed, it is only the light from a black body that represents by itself the brightness of the enclosure, and such a body, when taken out and hastily examined in the dark, without allowing it time to cool, will be found to give out rays having a brightness in all respects the same as that of the enclosure in which it was placed, because being opaque and non-reflective, all the light which it gave out in the enclosure was proper to itself, none having passed through its substance or been reflected from its surface; it therefore retains this light when taken into the dark, provided its temperature is not in the meantime allowed to fall*" [10, p. 277–278]. Experimental blackbodies of the 19th century were manufactured using either graphite, or soot [7], precisely because such carbon surfaces were not transparent and exceeded all others in being devoid of reflection.

In 1867, less than two years after Secchi and Faye [1] had conceived their solar model, G.J. Stoney explicitly placed carbon on the Sun: "*We have strong reasons for suspecting that the luminous clouds consists, like nearly all the sources of artificial light, of minutely divided carbon; and that the clouds themselves lie at a very short distance above the situation in which the heat is so fierce that carbon, in spite of its want of volatility, and of the enormous pressure to which it is there subjected, boils. The umbra of a spot seems never to form*

unless when the region in which carbon boils is carried upwards, or the hot region above the clouds is carried downwards, so as to bring them into contact, and thus entirely obliterate the intervening clouds..." [11]. Stoney's proposal introduced graphite particles in the photosphere, while reaffirming Faye's contention that the Sun was devoid of a distinct surface [1]. These words were to guide solar physics for two generations.

For instance, in 1891, during his Inaugural Address before the British Association, William Huggins stated: "*The Sun and stars are generally regarded as consisting of glowing vapours surrounded by a photosphere where condensation is taking place, the temperature of the photospheric layer from which the greater part of the radiation comes being constantly renewed from the hotter matter within... Consequently, we should probably not go far wrong, when the photosphere consists of liquid or solid particles, if we could compare select parts of the continuous spectrum between the stronger lines, or where they are fewest... The brightness of a star would be affected by the nature of the substance by which the light was chiefly emitted. In the laboratory, solid carbon exhibits the highest emissive power. A stellar stage in which radiation comes, to a large extent, from a photosphere of solid particles of this substance, would be favourable for great brilliancy... It may be that the substances condensed in the photosphere of different stars may differ in their emissive powers, but probably not to a great extent*" [12, p. 375–376].

Overall, the *Inaugural Address* amplified the search to understand the continuous nature of the solar spectrum. Huggins was a central figure in the history of solar astronomy and lived just prior to the conceptualization of a *fully* gaseous Sun. As such, it is almost as if his mind was suspended between two separate physical realities. He oscillated between a carbon containing photosphere as a source of light and a continuous spectrum produced exclusively by gases: "*We must not forget that the light from the heavenly bodies may consist of the combined radiations of different layers of gas at different temperatures, and possibly be further complicated to an unknown extent by the absorption of cooler portions of gas outside*" [12, p. 373]. The presentation by Huggins demonstrates a strained application of logic. Immediately after stating that: "*Experiments on the sodium spectrum were carried up to a pressure of forty atmospheres without producing any definite effect on the width of the lines which could be ascribed to the pressure. In a similar way the lines of the spectrum of water showed no signs of expansion up to twelve atmospheres; though more intense than at ordinary pressures, they remained narrow and clearly defined*" [12, p. 373], he writes: "*It follows, therefore, that a continuous spectrum cannot be considered, when taken alone, as a sure indication of matter in the liquid or the solid state*" [12, p. 373]. The experiments just described were contrary to the result sought. Ultimately, there could be no evidence that a gas could produce a blackbody spectrum simply by being pressurized. The spectrum may well have gained a continuous nature, but never with the proper blackbody shape. Huggins continued: "*Not only, as in the experiments already mentioned, such a spectrum may be due to gas when under pressure, but, as Maxwell pointed out, if the thickness of a medium, such as sodium vapor, which radiates and absorbs different kinds of light, be very great, and the temperature high, the light emitted will be of exactly the same composition as that emitted by lampblack at the same temperature, for the radiations which are feebly emitted will also be feebly absorbed, and can reach the surface from immense depths*" [12, p. 373]. In bringing forth these ideas from Maxwell, Huggins was abandoning the carbon containing photosphere.

James Maxwell wrote extensively about the theory of heat radiation [13]. He was well acquainted with Stewart and claimed: "*Professor Balfour Stewart's treatise contains all that is necessary to be known in order to make experiments on heat*" [13, p. vi]. In this regard, Maxwell's text contains many of the same ideas [13, p. 210–229] found in Stewart's works [14]. Maxwell's treatise also contained the classic lines previously invoked by Huggins [12, p. 373]: "*If the thickness of a medium, such as sodium-vapour, which radiates and absorbs definite kinds of light, be very great, the whole being at a high temperature, the light emitted will be exactly the same composition as that emitted from lampblack at the same temperature. For though some kinds of radiation are much more feebly emitted by the substance than others, these are also so feebly absorbed that they can reach the surface from immense depths, whereas the rays which are so copiously radiated are also so rapidly absorbed that it is only from places very near the surface that they can escape out of the medium. Hence both the depth and the density of an incandescent gas cause its radiation to assume more and more the character of a continuous spectrum*" [13, p. 226]. This conjecture, by Maxwell, was never validated in the laboratory. Sodium gas could not approach the blackbody spectrum under any circumstances, especially in the absence of a perfectly absorbing material. Even modern high pressure sodium lamps [15] could not produce the required spectrum. Their real emission was far from continuous and not at all like a blackbody [15, p. 23]. Nonetheless, Maxwell's theory became an anchor for those who believed that gases, if sufficiently thick, could produce a blackbody spectrum.

Astrophysics stood at an impasse between the need for carbon and its elimination from the solar body. Soon after Huggins delivered his famous address, William Wilson would approach the same subject in these words: "*Solar physicists have thought that the photosphere of the Sun consists of a layer of clouds formed of particles of solid carbon. As the temperature of these clouds is certainly not below 8000°C., it seems very difficult to explain how carbon can be boiling in the arc at 3500° and yet remain in the solid form in the Sun at 8000°. Pressure in the solar atmosphere seemed to be the most likely cause of this, and yet, from other physical rea-*

sons, this seemed not probable" [16]. Wilson goes on to state: "*carbon may exist in the solid form at very high temperatures although the pressures are comparatively low*" [16]. He was arguing in favor of solid carbon on the Sun despite the elevated temperatures. In 1897, along with George FitzGerald, Wilson would reaffirm his conviction while advancing an alternative for sunspots: "*Dr. Stoney called attention to an action of this kind that might be due to clouds of transparent material, like clouds of water on the Earth, but in view of the high solar temperature it seems improbable that any body, except perhaps carbon, could exist in any condition other than the gaseous state in the solar atmosphere; so that it seems more probable that Sun-spots are due, at least partly, to reflections by convection streams of gas, rather than by clouds of transparent solid or liquid particles*" [17].

Despite Huggins' *Inaugural Address*, Robert Ball, the Lowndean professor of astronomy and geometry at Cambridge, also reemphasized the central role of carbon in the structure of the Sun at the end of the 19th century: "*The buoyancy of carbon vapor is one of its most remarkable characteristics. Accordingly immense volumes of the carbon steam in the Sun soar at a higher level than do the vapors of the other elements. Thus carbon becomes a very large and important constituent of the more elevated regions of the solar atmosphere. We can understand what happens to these carbon vapors by the analogous case of the familiar clouds in our own skies... We can now understand what happens as the buoyant carbon vapors soar upwards through the Sun's atmosphere. They attain at last to an elevation where the fearful intensity of the solar heat has so far abated that, though nearly all other elements may still remain entirely gaseous, yet the exceptionally refractory carbon begins to return to the liquid state. At the first stage in this return, the carbon vapor conducts itself just as does the ascending watery vapor from the earth when about to be transformed into a visible cloud. Under the influence of a chill the carbon vapor collects into a myriad host of little beads of liquid. Each of these drops of liquid carbon in the glorious solar clouds has a temperature and a corresponding radiance vastly exceeding that with which the filament glows in the incandescent electric lamp. When we remember further that the entire surface of our luminary is coated with these clouds, every particle of which is thus intensely luminous, we need no longer wonder at that dazzling brilliance which, even across the awful gulf of ninety-three millions of miles, produces for us the indescribable glory of daylight*" [18].

The idea that the photosphere consisted of carbon containing luminous clouds would be echoed by almost every prominent astronomer of the 19th century, from Simon Newcomb [19, p. 269] to Charles Young [20, p. 194]. The finest spectroscopists, including John Landauer and John Bishop Tingle [21, p. 198–200], joined their ranks. Even in 1913, the ideas of Johnstone Stoney [11] were mentioned throughout much of professional astronomy, as reflected by the writings of Edward Walter Maunder [22]. Mauder, who had discovered the great minimum in the sunspot cycle, wrote about the solar constitution in these words: "*The Sun, then, is in an essentially gaseous condition, enclosed by the luminous shell which we term the photosphere. This shell Prof. C. A. Young and the majority of astronomers regard as consisting of a relatively thin layer of glowing clouds, justifying the quaint conceit of R. A. Proctor, who spoke of the Sun as a "Bubble"; that is a globe of gas surrounded by an envelope so thin in comparison as to be mere film. There has been much difference of opinion as to the substance forming these clouds, but the theory is still widely held which was first put forward by Dr. Johnstone Stoney in 1867, that they are due to the condensation of carbon, the most refractory of all known elements. Prof. Abbot, however, refuses to believe in a surface of this nature, holding that the temperature of the Sun is too high even at the surface to permit any such condensation*" [22].

Change was eminent and graphite was soon irrevocably cast out of the photosphere. In their 1885 classic text *On Spectrum Analysis*, Henry Roscoe and Arthur Schuster [23, p. 229–264] had already chosen to neglect the prevailing ideas relative to solar constitution. Arthur Schuster [24, 25] was soon to prepare his report on *Radiation through a Foggy Atmosphere* [26, 27]. With its publication, the decisive step towards the fully gaseous Sun would be taken and graphite soon forgotten.

2 The rise of theoretical astrophysics

Through Secchi and Faye [1], observational astronomers gazed upon a gaseous Sun. They could only dream of what they had created, as the concept of an ethereal star had evolved virtually in the complete absence of mathematical guidance. At the same time, though the photosphere maintained some semblance of condensed matter, the introduction of a tenuous solar interior provided a compelling invitation to theoretical study. If the Sun was truly a gas, then perhaps some understanding could be harnessed through the ideal gas law, which had been discovered by Clapeyron [28]. In contrast, William Herschel's solid Sun was devoid of such appeal [1]. The same was true true for Spencer's model. Though his solar interior was gaseous, his photosphere was liquid [1].

As for a fully gaseous Sun, the idea was full of theoretical promise. But was the interior of the Sun truly gaseous? For men of the late 19th and 20th century, there could be no question of this reality, in light of Andrews' discovery of critical temperatures [29]. Alfred Fisen would leave no doubt as to the importance of critical phenomena for solar models: "*The question as to the physical conditions existing in the interior of the Sun is attended with graver difficulty... When the necessity for the interior heat of the Sun being at least as high as that of its exterior became recognized, the solid globe was generally replaced by an ocean of molten matter. It is, however, scarcely possible to regard as existing in the interior of*

the Sun, matter in either the solid or in the liquid condition... It was for a time regarded as barely possible that the enormous pressure that must exist at great depths in the interior of the Sun might be effective in maintaining matter in the solid or liquid condition in spite of the high temperature, since it is a familiar fact in laboratory experience, that liquefaction of a gas is in every case assisted by pressure, and may in many instances apparently be affected by it alone. Since, however, it became apparent from the classic research of Dr. Andrews in 1869, that there exists for every element a critical temperature, above which it is impossible for it under any conditions of pressure to assume the liquid state, it has generally been regarded that the liquid interior to the Sun is next to an impossibility" [30, p. 36–37]. Armed with Andrews' discovery, the path seemed clear. Much of theoretical physics adopted a gaseous solar interior. They would eventually move forward to a *fully* gaseous structure, undaunted by the prospect that graphite or soot remained unchallenged as unique sources of blackbody spectra on Earth.

2.1 Friedrich Zöllner's protuberances: The laws of gases and the solar constitution

Zöllner was amongst the first scientists to apply the laws of gases to the study of the solar constitution [31, 32]. He attempted to understand the nature of solar protuberances, considering both eruptive flares and prominences. These works were important for two reasons: 1) Zöllner mathematically addressed the internal temperature for the Sun [33, 34] and 2) he highlighted that flares could not be easily explained when the Sun was considered *fully* gaseous. Using an atmospheric temperature of 27,700°C, Zöllner surmised that, at a depth lying 1/36th of the solar radius from the surface, the solar temperature approached 68,400°C [31].

Zöllner reasoned that eruptive protuberances, or solar flares, must occur because "*of a difference in pressure between the gases in the interior and those on the surface of the Sun*" [31]. In order to have an interior and an exterior, a boundary was certainly needed. Zöllner envisioned: "*Respecting the physical constitution of this layer, the further assumption is necessary that it is in some other state than the gaseous. It may be either solid or liquid. In consequence of the high temperature the solid state is excluded; and we must therefore conclude that the layer of division consists of an incandescent liquid*" [31]. Zöllner actually considered two models: "*Respecting the mass of hydrogen enclosed by this liquid layer, two suppositions appear to be possible*" [31]. The first was essentially a restatement of Spencer's "Bubble Sun" [35, 36] — a liquid photosphere with a gaseous interior [1]: "*The whole interior of the Sun is filled with glowing hydrogen, and our luminary would appear like a great bubble of hydrogen surrounded by an incandescent atmosphere*" [31]. At the same time, he considered a second situation in which the Sun was essentially liquid throughout while containing pockets of gas: "*The masses of hydrogen which are thrown out in these volcanic outburts are local aggregations contained in hollow spaces formed near the surface of an incandescent liquid mass, and these burst through their outer shell when the increased pressure of the materials in the interior reaches a certain point*" [31].

Zöllner would look back to Kirchhoff [37] and created a strange mix with the ideas of Secchi [38, 39] and Faye [40]. He placed the *fully* liquid layer, required in the interior of the Sun, at the level of the umbrae of sunspots [31, p. 319–320]: "*Hence it follows that the radius of the visible disk need not be necessarily identical with that of the supposed layer of separation, but that this latter may probably be assumed to lie below the point at which the hydrogen gas under compression evolves a continuous spectrum*" [31]. In doing so, Zöllner maintained the importance of the liquid layer in a manner completely independent of the need to generate the thermal spectrum. The enclosure provided by the liquid was required for the generation of flares. In fact, Zöllner argued against the need for condensed matter in producing the thermal spectrum: "*It is thus clear that it is not necessary, in order to explain the presence of dark lines in the solar spectrum, to assume that the continuous spectrum is produced by the incandescence of a solid or liquid body; for we may with equal right consider that the continuous spectrum is produced by the glowing of a powerful compressed gas*" [31]. By introducing this new layer, Zöllner advanced another reason why the Sun must possess condensed matter.

In treating the second scenario, that of a fully liquid Sun with pockets of gas, Zöllner made several arguments leading to a liquid solar interior: "*If we assume that the highest limit of specific gravity of this layer is the mean specific gravity of the Sun, we shall have to assume that all the deeper-lying layers, and therefore the sill deeper-lying gaseous layer, have the same temperature. But then the interior of the Sun would not consist of a gas, but of an incompressible liquid... In this case, however, the first supposition change into the second, according to which the Sun consists of an incompressible liquid...*" [31]. After completing several calculations, he then argued that pressures were rapidly increasing towards the solar interior. On this basis, the Leipzig professor rendered plausible the concept that the interior of the Sun could be liquid, despite high temperatures [31, p. 324].

In his second treatment on the solar constitution, Zöllner concentrated on determining the temperature of the chromosphere [32] and on refining the mathematical approach he had previously adopted. The 1873 article emphasized that line broadening could be affected by pressure, temperature, and optical thickness of the sample [32]. In this regard, Zöllner was concerned with the quantity of luminous particles in the line of sight of the observer. As such, he elucidated the complex considerations involved in obtaining temperatures and densities from the line widths of gases near the solar surface. Zöllner's second treatise was devoid of the complex solar theories which had characterized his first work [31].

2.2 Jonathan Homer Lane: A gaseous Sun endowed with condensed matter

In his *Memoire*, Cleveland Abbe presented a detailed picture of J. Homer Lane [41]. Lane considered Helmholtz's theory and Espy's theory of storms, while applying the ideal gas law to the Sun [42]. In so doing, he became the first scientist to build a truly mathematical model of a gaseous star. Like Einstein, Lane had worked as a patent examiner. He was said to have been quiet and lacking the fluency of speech [41]. Lane was never married and he was personally known to only a few people [41, p. 259]. He was deeply religious and he displayed many marks of simple nobility. Cleveland recounted these in the words of Byron Sunderland: "*Of the propriety, integrity, and simplicity of his life, of his exceeding conscientiousness and carefulness and his modest shrinking from all self-assertion or ostentation, we all well know. He was not what we should style a demonstrative man. He lived quietly within himself, and his life was engrossed in scientific pursuits. The nature and construction of his mind was purely mathematical. This was evident in the exactitude of his language, even in the most casual conversations and the most trivial subjects*" [41, p. 261].

Stevenson-Powell provided a detailed and extensive review of Lane's classic work on the theoretical modeling of a gaseous Sun [43]. In his approach to science, Lane was not unlike Eddington [44] and chose to consider the Sun as a theoretical physicist. He proposed a model and then considered the ramifications [43, p. 190], tackling a question by extrapolating from the known laws of physics. At the same time, "*Lane had little interest in the physical appearance of the Sun, and none at all in the spectral discoveries that increasingly influenced ideas about the Sun during the 1860s*" [43, p. 183]. The same could be said of Eddington [44].

Lane was responsible for advancing the first of the polytropic gas spheres. He was followed in this endeavor primarily by August Ritter [45], William Thomson (Lord Kelvin) [46], and Robert Emden [47]. Subrahmanyan Chandrasekhar provided a detailed treatment of polytropes in his classic text *An Introduction to the Study of Stellar Structure* [48, p. 84–182] whose bibliographical notes included excellent summaries of all key contributions in this subject area. Eddington also discussed the polytropes in *The Internal Constitution of the Stars* [44, p. 79–96].

Lane based his theoretical contribution on the ideal gas and Espy's theory of storms, advanced more than twenty years earlier [42]. But, the concept that the Sun was an ideal gas created obstacles. Stevenson-Powell recounted this fact, citing Arthur Eddington: "*In Lane's time there was no evidence that any star existed for which the theory of a perfect gas would be applicable*" [43, p. 190]. While the work of Andrews on critical temperatures was already well recognized [29], many failed to completely abandon the idea that the Sun contained at least some condensed matter.

In spite of these difficulties, the American scientist viewed the Sun as a gaseous sphere possessing a condensed exterior. He opened his classic paper as follows: "*Some years ago the question occurred to me in connection with this theory of Helmholtz whether the entire mass of the Sun might not be a mixture of transparent gases, and whether Herschel's clouds might not arise from the precipitation of some of these gases, say carbon, near the surface, with the revaporization when fallen or carried into the hotter subjacent layers of atmosphere beneath; the circulation necessary for the play of this Espian theory being of course maintained by the constant disturbance of equilibrium due to the loss of heat by radiation of the precipitated clouds*" [42]. Lane was replaying the ideas of Stoney, Secchi, and Faye [11, 38–40]. Nonetheless, the study of Lane's private notes revealed an unpublished paper from 1867 *The Sun viewed as a gaseous body* [43, p. 186]. In these unpublished notes, Lane claimed priority of ideas and wrote: "*The within formulae were written down about the year 1863 (perhaps earlier) considering the credibility of the Sun being a gaseous body, sustaining its heat by the descent of its mass in cooling, and keeping up by its circulation a continual precipitation of (carbon?) vapor in the photosphere, and the continual re-vaporization of the carbon? in the interior, after the philosophy of terrestrial storms as explained by Espy. Conclusion: it seemed evident the Sun's gaseous constitution could not be credibly referred to the laws of the gases, so far as they are known. J.H.L. May 1867*" [43, p. 187]. It appeared that Lane might have conceived of a gaseous Sun independently, in 1863. However, it would be difficult to conceive that such similarity with the well-known works of Secchi and Faye was purely coincidental [38–40]. Lane properly claimed that Faye's theory was "*seriously lacking*" [42]. The 1865 articles, by the French author, were devoid of mathematical treatment [1]. Through Lane's work, carbon was once again mentioned. Hence, even in the first truly theoretical work on a gaseous Sun [42], the emissivity of graphite maintained its powerful undercurrent.

2.2.1 Lane and convective equilibrium

Interestingly, Lane used the concept of convective equilibrium as a footnote to his first equation [42]. William Thomson had proposed the existence of convective equilibrium in 1862 and applied the idea to a gaseous Sun in 1887 [46]. By this time, Lord Kelvin had abandoned his original idea that the Sun was liquid [1]. Convective equilibrium would become one of the great building blocks of the theory of a gaseous Sun. Chandrashekhar would cite Kelvin's understanding of convective equilibrium in his classic text [48, p. 85]: "*If a gas is enclosed in a rigid shell impermeable to heat and left to itself for a sufficiently long time, it settles into the condition of gross-thermal equilibrium by 'conduction of heat' till the temperature becomes uniform throughout. But if it were stirred artificially all through its volume, currents not considerably disturbing the static distribution of pressure and*

density will bring it approximately to what I have called convective equilibrium of temperature. The natural stirring produced in a great fluid mass like the Sun's by the cooling at the surface, must, I believe, maintain a somewhat close approximation to convective equilibrium throughout the whole mass" [46].

Convective equilibrium was a strange allusion, given that convection, by definition, was a non-equilibrium process. Convection existed as a result of the second law of thermodynamics, a principle first outlined by Clausius [49, 50] and ironically, by William Thomson [51]. To call for convective equilibrium "artificially" implied a violation of the first law of thermodynamics. To invoke it on the Sun, was a violation of the second law. Convective equilibrium could never exist, either on or within the Sun precisely because, by its very nature, convection was a non-equilibrium process. True system equilibrium required that both conduction and convection be absent. In Lane's case, recourse to convective equilibrium for his mathematics was particularly unusual, given that he had opened his manuscript with the statement that: "*the circulation necessary for the play of this Espian theory being of course maintained by the constant disturbance of equilibrium due to the loss of heat by radiation of the precipitated clouds*" [42]. How could a theory of storms ever form the basis for invoking convective equilibrium?

2.2.2 Lane and the temperature of the solar surface

The final portion of Lane's paper centered on elucidating the temperature at the upper visible solar surface. He reached the conclusion that this number must not be too far from 54,000°F and raised an objection to Faye's model: "*It must be here recollected that we are discussing the question of clouds of solid or at least fluid particles floating in a non-radiant gas, and constituting the Sun's photosphere. If the amount of radiation would lead us to limit the temperature of such clouds of solids or fluids, so also it seems difficult to credit the existence in the solid or fluid form, at a higher temperature than 54,000° Fah. of any substance that we know of*" [42].

Though Lane adopted Faye's model as a point of departure, he was open, though non-committal, to the idea that the Sun was fully gaseous: "*Dr. Craig, in an unpublished paper, following the hint thrown out by Frankland, is disposed to favor the idea that the Sun's radiation may be the radiation of hot gases instead of clouds. At present, I shall offer no opinion on that point one way or another, but will only state it as my impression that if the theory of precipitated clouds, as above presented, is the true one, something quite unlike our present experimental knowledge, or at least much beyond it, is needed to make it intelligible*" [42]. Craig was referring to the classic paper by Lockyer and Frankland discussed in Part I of this work [1]. Clearly, Lane had strong reservations relative to Faye's model, even though it formed the basis for much of his own presentation.

Lane advanced two ideas to uphold the precipitated cloud theory. In the first, he invoked Clausius' work on the specific heat of gases, using the idea that hydrogen might be able to exist, either in atomic or molecular form [42]. This was a novel concept at the time and Lane believed that the precipitated cloud model could be preserved through its introduction. However, the most fascinating defense was found in his second hypothesis which he believed was not very sound and dismissable with very little reflection [42]. Interestingly, in this hypothesis, Lane abandoned varying densities in the solar interior and created the requirements for a liquid Sun, apparently without realizing the obvious change in phase and the profoundness of his own writings. Lane advanced the possibility that "*in the Sun's body the average length of the excursion made by each molecule between two consecutive collisions, becomes very short compared to the radius of the sphere of repulsion of molecule for molecule, and with the average distance of their centers at nearest approach. This way of harmonizing the actual volume of the Sun with a temperature of 54,000° Fah. in the photosphere, and with the smallest density which we can credit the photosphere, would involve the consequence that the existing density of almost the entire mass of the Sun is very nearly uniform and at its maximum possible, or at all events that any further sensible amount, comparatively, of renewed supplies of heat, for the obvious reason that this hypothesis carries with it almost the entire neutralization of the force of gravity by the force of molecular repulsion*" [42]. Lane, without direct reference, was calling for a liquid Sun. He concluded: "*Another thing involved in this second hypothesis is the fact which Prof. Peirce has pointed out to the Academy, viz: that the existing molecular repulsion in the Sun's body would immensely exceed such as would be indicated by the modulus of elasticity of any form of matter known to us*" [42]. With these words, Lane reminded his readers that the conditions within the Sun were very different than those predictable at the time using terrestrial physics. Given the pressures within the Sun, the possibility of unusual materials had to be considered. For Lane, this extended to a material approaching a liquid in behavior, even though such conjectures were viewed as unlikely.

2.2.3 Lane's law: Stars which cannot cool

In his 1870 treatment of the Sun [42], Lane advanced an elegant approach to the gaseous Sun. From his mathematics, he was able to obtain a relationship between solar density and radial position using two equilibrium conditions. Today, these are referred to as 1) mechanical or hydrostatic equilibrium and 2) convective equilibrium. At the same time, Lane deduced a central solar density of 7 to 28 g/cm^3 depending on the assumptions applied [42]. Yet, the most important conclusion of Lane's paper was a law, not discovered by Lane but by Ritter [45]. In fact, Chandrashekhar would state that "*almost the entire foundation for the mathematical theory of stellar structure was laid*" by Ritter [48, p. 179].

As for Lane's law, it proposed that the product of a gas-

eous star's radius and its radial temperature was a constant [43, p. 194]. If the star contracted, its temperature increased, provided that it remained an ideal gas. Fisen commented as follows: "*In a very remarkable paper, published in 1870, Mr. Homer Lane has shown that if the Sun were entirely gaseous, and if the gases composing it were under such physical conditions that the laws of 'perfect gases' should be applicable to them, the heat developed by shrinkage must not be merely equal but must so far exceed that radiated to effect it, that the temperature of the whole must actually rise in consequence, and must continue to do so for so long as a perfectly gaseous condition is maintained*" [30, p. 38]. Professor Benjamin Peirce would restate the same ideas: "*Gaseous bodies in the process of radiating light and heat condense and become hotter throughout their mass*" [52, p. 197–198]. Today, "*Lane's Law*" is referred to as Lane-Emden equation, even though Ritter discovered the formula and Lane never wrote it down [43, p. 196]. As a result of the Lane-Emden relation, gaseous stars could never cool. They continued to emit massive amounts of heat radiation. In so doing, gaseous stars actually contracted and heated up. Eddington was astounded at the "*striking result that if a star contracts the internal temperature rises so long as the material is sufficiently diffuse to behave as a perfect gas*" [44, p. 5].

2.2.4 An independent discovery of Lane's law

Lane's law was also independently discovered by T.F.F. See [53, 54]. See provided a detailed description of his experiences with Lane's law. The discourse was both credible and instructive [54]. See's treatment of Lane's law advanced a straightforward derivation from Helmholtz' ideas and placed much of the history of Lane's law in perspective. Ritter's work was not very well known by the astronomical community. After deriving Lane's law, See recognized its profound importance and wrote to many astronomers to establish if there were priority claims to the formulation. Eventually, an English astronomer mentioned Ritter's 1881 communication [54]. Examining the reference, See argued that Ritter only used "language" to describe Lane's law. In fact, as Chandrashekhar stated [48, p. 178], Ritter first arrived at the law in the key 1878 paper [45]. Unfortunately for See, the Englishman was poorly aware of the German literature. In large measure, See's own work, would simply become an independent confirmation of Ritter.

However, See's papers were both elegant and well written [53, 54]. See argued that star-like masses, formed from nebular bodies, could not become infinitely compressed. Eventually, they must reach the liquid state: "*From these considerations we see that when the gaseous nebula is infinitely expanded the temperature is the absolute zero of space, and that the maximum temperature results when the mass is contracted to the smallest radius consistent with the laws of gaseous constitution. After the mass has condensed so far that liquefaction sets in, free contraction is obstructed by molecular forces, or practically ceases; the temperature falls, and the body eventually cools down to obscurity. Such it would seem, must be the history of the temperature of cosmical bodies formed by the gravitational condensation of nebulous matter*" [54]. For theoretical astrophysics, it was difficult to account for such a phase transition.

2.3 Charles Hastings: A photosphere made of silicon?

When Charles Hastings developed his theory on the constitution of the Sun, he was surely unaware of the great impact he would have on solar theory [55]. Though Hastings' contribution was devoid of mathematics, it advanced many novel ideas which became the genesis for new theoretical formulations. Amongst his contributions was the concept that line widths could be explained by considering various layers within the photospheric atmosphere. For Hastings, line widths were directly related to pressure [55]. In order to arrive at increasing values, it was simply required that the lines originated from deeper layers within the photosphere.

Hastings opposed Faye's model of the Sun on two grounds: "*1) To produce dark lines in a spectrum by absorption, the source of the absorbed light must be at a higher temperature than that of the absorbing medium and 2) There is an inferior limit of brightness below which the course of absorbed light cannot go without the spectral lines becoming bright*" [55]. In the second of these objections, Hastings was referring to the reversing layer of the Sun observed during total eclipses.

Hastings advocated that "*it is not a priori improbable that we receive light from many hundreds of miles below the outer surface of the photosphere*" [55], a concept still utilized in the modern age to explain limb darkening. Hasting applied the idea to explain the linewidths of dark lines in the solar spectrum and proposed an alternative approach to account for limb darkening. Hastings also advocated that solid or liquid carbon could not be present on the Sun: "*Granting this, we perceive that the photosphere contains solid or liquid particles hotter than carbon vapor, and consequently not carbon*" [55]. He suggested that the material might be silicon. Hastings made the bold pronouncement: "*At any rate, we are sure that the substance in question, so far as we know it, has properties similar to those of the carbon group*" [55]. But what properties? Hastings was not clear on this point. Nonetheless, the idea was important and Hastings' point will be addressed in an upcoming contribution [56].

2.4 Frank Very: Frequency dependent limb darkening

In 1902, Frank Very published a detailed analysis of limb darkening as a function of frequency [57]. The work would be monumental in astronomy. Very was once Samuel Langley's trusted assistant [58] and had been with Langley in the days when the solar spectrum was first recorded in its entirety [59–61]. In his classic report [57], Very documented

that the Sun's radiance was darkening towards the limb in a frequency dependent manner. He studied 7 wavelengths ranging from 0.416 μm to 1.5 μm, and demonstrated that shorter wavelengths produced more dramatic limb darkening [57]. In the violet wavelengths (0.416 μm), the edge of the solar disk was radiating only 10% of the intensity found at the center. As one moved towards the red (1.50 μm), the decrease was much smaller with 75% of the radiation remaining [57].

Very attempted to explain his findings by invoking atmospheric absorption of radiation, primarily by the corona [57, p. 80]. Very advanced the scattering of radiation in the corona and its reflection by carbon particles [57, p. 82]. Of course, graphite makes for a very poor reflector. Very considered diffraction: "*We can subject the hypothesis of an extensive envelope, depleting the rays by selective diffraction*" [57]. Finally, Very advanced that the phenomenon was produced by the irregularity of the Sun's photosphere, invoking its granulated structure [57, p. 86]. The idea was never pursued.

Immediately following the publication of Very's discovery, Arthur Schuster attempted to explain the strange frequency/position dependent variation of solar radiation [62]. In so doing, he began to develop the logic which led to his famed communication on *Radiation through a Foggy Atmosphere* [26, 27]. Very's work became a source of motivation for theoretical physics.

2.5 Arthur Schuster and the solar atmosphere

Sir Arthur Schuster was one of the most influential scientists of his time [24, 25]. He attended Balfour Stewart's classes and, following the counsel of Henry Roscoe, completed his dissertation with Gustav Kirchhoff [24, 25]. At the Cavendish Laboratory, Schuster worked under both James Clerk Maxwell and Lord Rayleigh [24]. He also studied with Weber and Helmholtz [25]. In 1888, he succeeded Balfour Stewart as the Langworthy Professor of Physics at Owen's College and remained in this chair until 1907 [25]. Eventually, Schuster was elected secretary of the Royal Society [24]. If George Hale was regarded as the "*father*" of the International Union for Solar Research, it has been argued that Schuster was its "*mother*" [25]. Schuster counted amongst his students Sir J. J. Thomson (Nobel Prize 1906), John William Strutt (Lord Rutherford, Nobel Prize 1904), and Sir Arthur Eddington, [24]. As a consequence, Eddington became a direct scientific descendent of Gustav Kirchhoff.

Schuster's seminal contributions began in 1902 with a report on *The Solar Atmosphere* published within the *Astrophysical Journal* [62]. *The Solar Atmosphere* was written in response to Frank Very's detailed examination of solar radiation [57] (see Section 2.4). In turn, it was subjected to a letter of criticism authored by Very [63] to which Schuster would reply [64].

Schuster's reply, *The Temperature of the Solar Atmosphere* [64], summarized his position and exposed some rather prominent errors in logic. Schuster believed that he could account for the law of variation of solar radiation by invoking two layers within the Sun: 1) a photospheric layer radiating as a blackbody at 6,700° and 2) an absorbing layer at 5450°. The sum of the two layers produced the Sun's apparent temperature at 6,000°. Schuster stated that within *The Solar Atmosphere* [62], he had used a fourth power of temperature relationship, when a fifth power was more appropriate. Additionally, and this was perhaps most troubling, Schuster maintained that the radiative layer was emitting as kF, where F was the blackbody function and k was a wavelength dependent constant which could adopt any value between zero and infinity. In so doing, he removed all restrictions on the ability of bodies to emit radiation and operated well outside the bounds of physics. As a student of Kirchhoff, Schuster insisted that: "*Everybody knows that the function of temperature and wavelength which expresses the radiation of a blackbody is a fundamental function which must enter into every discussion of radiation and absorption*" [64]. Yet, through his mathematics, Schuster essentially disregarded the blackbody function itself. Schuster could provide no physical justification for the behavior of k, his magical constant. Its presence made any extended discussion of mathematics pointless. Schuster further broadened the boundary of proper mathematical treatment highlighting: "*As misunderstandings seem so easily to arise, it is perhaps worth pointing out that, although for the purpose of facilitating mathematical analysis it is sometimes necessary to treat the upper portion of the same body as made up of distinct layers, having different temperatures and possibly different absorbing qualities...*" [64]. With these words, Schuster removed even more restrictions for the gaseous solar models relative to ability to emit radiation. Given unbridled mathematics, all could be explained in a gaseous framework.

Very seemed more mindful of physical realities: "*It is a fact that, at the photospheric level, some form of matter exists which does radiate indiscriminately through a wide range of wavelengths, and whose particles are presumably coarse enough to act non-selectively in other respects*" [63]. He championed an idea that was to permeate theoretical astrophysics: "*From the depths of the Sun, radiations composed mainly of very short waves tend to proceed, and a very extensive scattering atmosphere acts almost like a reflector, send nearly all the rays back again. In this case the medium will not be heated much in the process. Only a small fraction of the incident rays will be absorbed by the fine particles; the greater part is assumed diffracted. Still, as the course of the rays through such an extensive scattering medium is a zigzag one, the scattering being repeated over and over again, some cumulative action and some absorption of energy by the medium must result. Consequently, it is not possible to separate completely the two causes — absorption and scattering*" [63]. Almost the exact arguments would be repeated by Eddington in the 1920's [44].

2.6 Classic papers in stellar radiation transfer

Donald Menzel prepared a compilation of *Selected Papers on the Transfer of Radiation* [66], wherein he reprinted the great contributions on the subject, but regrettably, without offering a commentary. By assembling these articles in one text, Menzel implicitly reminded the reader of their importance in the history of theoretical astrophysics.

The study of radiation in stellar atmospheres was primarily driven by the need to explain the continuous solar spectrum. While many works describe the transfer of radiation within stars [67–69], the entire problem was introduced into astronomy by the desire to account for thermal emission in a gaseous framework. The understanding of internal stellar opacity was directly associated with the act of building a star without recourse to condensed matter. Ironically, it also became essential to account for physical structure using a phase of matter, which on Earth, was devoid of structural potential [70]. In adopting a gaseous foundation, astrophysics was immediately confronted with two dilemmas: 1) how could a gas provide a continuous blackbody spectrum like graphite? and 2) how could structure and activity, like granulations, sunspots, flares, and prominences be understood using a fully gaseous entity? To solve these great questions, only theoretical approaches were available.

2.6.1 Schuster and the foggy atmosphere

Arthur Schuster initially presented an abridged version of his *Radiation through a Foggy Atmosphere* in 1903 [27]. The complete paper appeared in 1905 [26]. Schuster attempted to explain the bright lines of the reversing layer above the photosphere and the dark lines which usually typify the solar spectrum. For Kirchhoff, the bright lines were being produced by species which were at a higher temperature than the liquid photosphere, while the dark lines required lower temperatures. Though Kirchhoff's student, the German-born British physicist preferred an alternative explanation.

Schuster viewed as *foggy* an atmosphere which sustained a considerable amount of scattering. The basis of the presentation was the emission of radiation from a surface towards an overlaying atmospheric layer, wherein both scattering and absorption occurred. Accordingly, Schuster required that the Sun possess a distinct surface [26]. The point was also made by Milne [70] in his description of Schuster's contribution to the understanding of solar emission. For Schuster, scattering and absorption within the foggy atmosphere could modify the light emitted from the lower surface, permitting only certain frequencies to pass through which accounted for the bright or dark lines on the solar spectrum. The derivation assumed that the coefficient of absorption in the scattering layer was a function of wavelength dependent on the density of the absorbing species in the medium. Likewise, the coefficient of scattering also depended on the number of scattering particles in the medium which may or may not be the same as those used in absorption.

Schuster considered the Sun much like Faye [2]. The photosphere was composed of particulate matter floating above a gaseous solar body [1]. It was this particulate matter which would allow for the treatment of the scattering process. Schuster insisted on the validity of Kirchhoff's law as the proper starting point for all work in thermal emission. Though he recognized many of the weaknesses of his approach, Schuster never questioned Kirchhoff [26, p. 5]. Consequently, Schuster demonstrated that when the absorption coefficient of the layer was large with respect to the coefficient of scattering, the radiation observed from a large cloud of gas was the blackbody function: "*The radiation in this case becomes equal to that of a completely black surface, which agrees with the well-known law that absorption irrespective of scattering tends to make the radiation of all bodies equal to that of a black body when the thickness is increased*" [26, p. 6]. The result unfortunately, while mathematically appealing, was logically flawed.

Schuster expressed that the radiation emitted by the absorbing layer was the product of the absorption coefficient, k, multiplied by the blackbody function, E, and the thickness of the layer, dx: $kEdx$ [26, p. 3]. The absorption coefficient, k, in this case, was dependent on the wavelength of observation, the nature of the gas, and the density of the medium. In reality, Schuster needed to use an arbitrary function, like Γ, obtaining $k\Gamma dx$. In this case, Γ could be viewed as equal to $k'E$. Such an approach would more appropriately reflect the complexity involved in this problem. Schuster never established that E equaled Γ, the step critical to maintaining his conclusion. His *a priori* invocation of the blackbody function for the gas layer, though appearing mathematically correct because of the multiplication with k, ensured the result sought. Repeating the same derivation using Γ would completely alter the conclusions.

Once Schuster assumed that the blackbody function could be directly applied to represent the emission of the gas, a great thickness guaranteed that blackbody radiation was produced, even if the coefficient of absorption was small, merely because the coefficient of scattering was much smaller (see Eq. 14 in [26]). The result was impossible as it violated the first law of thermodynamics. It would have been more reasonable to derive that great thickness would simply result in obtaining the arbitrary function Γ. Schuster would have obtained this tempered finding, reminiscent of the line spectrum, such as that of the gaseous nebula in Orion [71, p. 87], if he had not insisted upon using the blackbody function as a point of departure.

The lineshapes of emission spectra for condensed matter do not change simply because objects become large. Yet, this was what Schuster was implying for the gas. This conclusion was very far reaching and would propagate throughout the astrophysical literature without correction. Arbitrary radiation never becomes black within adiabatic enclosures [72] and

gases do not become black simply because they are expansive — a lesson learned from the gaseous nebula [71, p. 81–92]. The size of objects remains secondary to the nature of radiation, if diffraction effects can be neglected [73].

2.6.2 Schwarzschild and radiative equilibrium

As was seen in Section 2.2, Lane's gaseous Sun [42] achieved stability through convective equilibrium. But for Arthur Eddington, radiative equilibrium became an important means of achieving the same result [3, 44]. The concept of radiative equilibrium was initially advanced, as Eddington recalls [44, p. 9], by R. A. Sampson in 1894 [74]. Still, it was Karl Schwarzschild (October 9th, 1873 — May 11, 1916) [75] who, in 1906, would give it prominence in theoretical astrophysics [76].

Schwarzschild was a gifted theoretical physicist who died at the age of 42 in the course of World War I: "*The war exacts its heavy roll of human life, and science is not spared. On our side we have not forgotten the loss of the physicist Moseley, at the threshold of a great career; now from the enemy, comes news of the death of Schwarzschild in the prime of his powers. His end is a sad story of long suffering from a terrible illness contracted in the field, borne with great courage and patience. The world loses an astronomer of exceptional genius, who was one of the leaders in recent advances both in observational methods and theoretical researches*" [75]. Many surely believe in the impossibility of reading Schwarzschild without gaining some reverence for the beauty of the human mind. Schwarzschild's treatment of radiative equilibrium within stars would not set a lower standard [76].

Milne reviewed Schwarzschild's contribution to radiative equilibrium in his Bakerian lecture [70]. This elegant treatment, as mentioned in Section 2.6.1, also addressed Schuster's approach [70].

Schwarzschild began his discussion of limb darkening on the solar surface by assuming that radiative equilibrium existed [76]. He also considered adiabatic equilibrium, referred to by Lane as convective equilibrium [42]. According to Schwarzschild: "*radiative equilibrium in a strongly radiating and absorbing atmosphere will be established when radiative heat transfer predominates over heat transfer due to convective mixing*" [76]. The theoretical formulation adopted resembled Schuster's [70]. Schwarzschild almost perfectly accounted for limb darkening using radiative equilibrium, demonstrating accordingly, that this assumption was valid for a gaseous Sun. The final result was independent of wavelength, dealing only with the total heat emitted, as measured with a bolometer [76]. Schwarzschild further proved that limb darkening could not be accounted for using convective equilibrium (see the table in [76]). The finding was impressive. Like Schuster before him, Schwarzschild based his conclusion on the validity of Kirchhoff's law [9]. Thus, the result was critically dependent on the soundness of Kirchhoff's conclusion. In addition, since it was based on an ideal gas, Schwarzschild's derivation implied that the Sun was devoid of a real surface and the solutions obtained extended to infinity [76]. Radiative equilibrium, sustained within a gaseous Sun, would form the basis of Eddington's treatment of the internal constitution of the stars [3, 44, 77].

2.6.3 Rosseland and mean opacities

Before discussing Eddington's application of radiative equilibrium to the stars, a sidestep should normally be made in order to briefly cover Rosseland and the formulation of the mean opacities [78, 79]. First proposed in 1924, Rosseland mean opacities enabled the next great advance in theoretical astrophysics [78, 79]. However, the topic will be passed over for the time being, reserving it instead for an upcoming work [80].

3 Eddington and Jeans: The clash of the titans

In writing the biography of Arthur Stanley Eddington, Subrahmanyan Chandrasekhar chose the following title: *Eddington: The Most Distinguished Astrophysicist of his Time* [81]. Chandrasekhar was not far from the mark. However, another contender for the title existed: James Hopwood Jeans. In fact, Edward Arthur Milne [82], who along with Ralph Fowler [83] worked with Eddington at Cambridge, would spend the last days of his life writing the biography of Sir James Jeans [84]. The work would be published after Milne's death. No one can truly dissect the merits of each man. Eddington and Jeans were giants in the world of theoretical astrophysics. Each made brilliant strides and, like all men, each committed regrettable scientific errors.

Matthew Stanley provided an outstanding account of the great battle which engulfed Eddington and Jeans [85]. Stanley outlined the vivid debates over the nature of the stars and the vastly differing philosophical approaches. He emphasized that much of what theoretical astrophysics would become dependent on Eddington's phenomenological outlook [85]. Jeans, for his part, dismissed Eddington's approach as not even science [85]: "*Eddington argued that his phenomenological approach opened up new avenues of investigation in astronomy, but Jeans argued that this was a violation of the very rigor and discipline that made astronomy so powerful*" [85]. Albert Einstein shared in Jeans' position stating: "*Eddington made many ingenious suggestions, but I have not followed them all up. I find that he was as a rule curiously uncritical towards his own ideas. He had little feeling for the need for a theoretical construction to be logically very simple if it is to have any prospect of being true*" [86, p. 40]. Einstein wrote these words in a private letter and made no such statements publicly. After all, it was Eddington who first worked to confirm Einstein's theory of relativity [87]. Jeans was even more critical: "*All Eddington's theoretical investigations have been based on assumptions which are outside the laws of physics*" [88]. As for Eddington, he was described

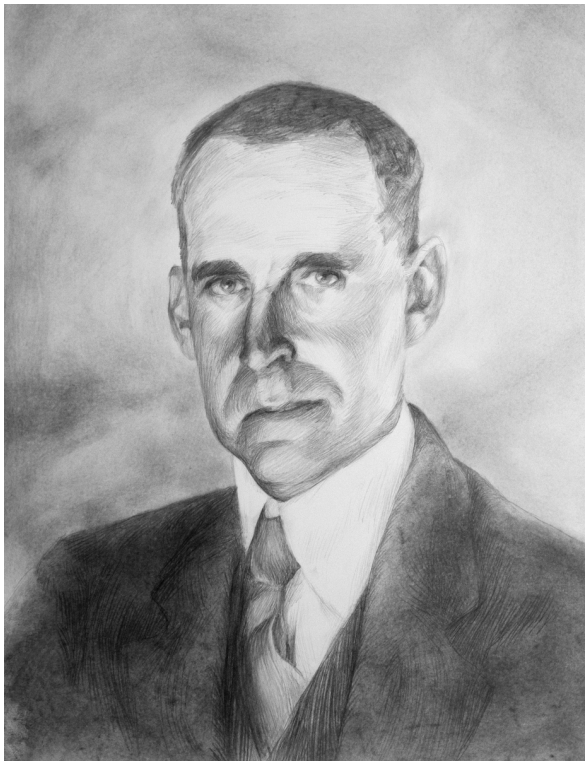


Fig. 1: Sir Arthur Stanley Eddington (December 28th, 1882 — November 22nd, 1944) was an outstanding theoretical physicist. He would become known for his approach to the gaseous stars. He derived a mathematical formulation which could account for the mass-luminosity relationship of the stars and was the first to propose that stars were fueled by nuclear processes. Eddington also conducted key experiments validating Einstein's theory of relativity.

Fig. 2: Sir James Hopwood Jeans (September 11th, 1877 — September 16th, 1946) was the last modern advocate of liquid stars. He believed that such objects were constructed from heavy elements obtaining their energy through fission, rather than fusion. Beyond astronomy, he was best known for his work on the partition of energy between matter and radiation — a solution leading to the Jeans-Rayleigh ultraviolet catastrophe. Jeans served as Secretary of the Royal Society from 1919–1929.

as a pragmatist [85]. He used "*whatever knowledge and tools were useful, instead of worrying about whether they were 'really true'*" [85]. In his defense against Jeans' constant detractions, Eddington claimed: "*although a reasonable degree of rigour is required, the laborious exploration and closing of every loophole is of secondary importance* [85]. But, with regards to the Sun, who was to assess if an element of theory was merely a question of closing a loophole or a fatal and irrecoverable logical flaw? Eddington and Jeans would outline scientific and philosophical problems which remained unanswered to the present day.

Milne, perhaps better than anyone, was in a position to highlight the great loss to science that the discord between Jeans and Eddington produced: "*It is much to be regretted that these two titans, Eddington and Jeans, should not have co-operated in their assaults on the grand subject of stellar structure, instead of being opposed to one another, during the most constructive periods of their careers. The blame has to be divided between them. Jeans mistakenly attacked Eddington's mathematics instead of accepting his mathematics and then providing the correct interpretation; Eddington resented what he considered to be aspirations on his competency as a mathematician, and never understood the difficulties of a philosophical kind that surrounded his own interpretation of his results. Astronomers on the whole have favoured Eddington's side of the controversy — mistakenly in my opinion. This is due, in addition to the reasons mentioned above, to the fact that Eddington had more of a feeling for the physics of a situation than Jeans had, whilst Jeans had more of a feeling for the mathematics of a situation than Eddington had; the result was that Eddington's stars had a physical plausibility that Jeans' lacked, and the astronomer who did not wish to go into the rights and wrongs of the mathematical situation could see the physical likelihood of Eddington being correct*" [84, p. 28].

3.1 Arthur Stanley Eddington

Though Eddington was a great proponent of the gaseous Sun, in 1910, he noted that "*the stars might be solid, liquid, or not too rare a gas*" [85]. He was a Quaker by birth and had earned a bachelor's degree with Arthur Schuster at Owens College [85]. As such, he was a direct scientific descendent of Kirchhoff. Eddington maintained that the value of theory

was in its ability to prompt further study, not in its relation to the established facts [85].

In his classic paper *Radiative Equilibrium in the Sun and Stars*, Eddington wrote about the laws of emission: "*There are some physical laws so fundamental that we need not hesitate to apply them to the most extreme conditions; for instance, the density of radiation varies as the fourth power of temperature, the emissive and absorbing power of a substance are equal, the pressure of a gas of given density varies as its temperature, the radiation-pressure is determined by the conservation of momentum — these provide a solid foundation for discussion*" [77]. Unfortunately, Eddington dispensed with the qualifiers so critical to make such statements hold true. In reality, only the emission of graphite or soot varied as the fourth power of temperature [7, 72, 73, 89]. Even for these cases, the relationship depended on the frequency of interest and the specific mineralogical origin of the material. The gas Eddington considered could never adopt such behavior [89]. In fact, the emissivity of gases could actually drop with increasing temperature [89], a clear violation of Stefan's 4th power of temperature law [90]. Unlike graphite, gases utilize convection currents in an attempt to reach thermal equilibrium. In any event, Kirchhoff's law [9] required two restrictions: a rigid enclosure and thermal equilibrium [7, 72, 73]. Eddington's gaseous Sun could provide neither. Outside the strict confines of thermal equilibrium, even the statement that emission equaled absorption was invalid. Jeans also made the point: "*In a gaseous star it is probable that much more energy is transferred by radiation than by ordinary gaseous conduction, so that an accurate determination of the laws of radiative transfer is a necessary preliminary to many problems in stellar physics*" [91]. Jeans based his thesis on theoretical grounds, while the laws of radiation for gases must be determined experimentally. In any case, even the slightest conduction and/or convection, both of which are undeniably present in stars, rendered all conjectures of radiative equilibrium invalid.

Despite all these considerations, Eddington was able to make what appeared to be surprisingly powerful advances in theoretical astrophysics. While assuming that absorption was constant within stars, the triumph of his gaseous models rested on the confirmation of the mass-luminosity relationship [44, p. 145–179] and the explanation of Cepheid variables [44, p. 180–215]. Eddington's paper, *On the Relation between the Masses and Luminosities of the Stars*, became an instant classic in theoretical astrophysics [92]. Eddington justified his theoretical approaches by invoking the work of Jacob Halm [93] who was the first to state that "*intrinsic brightness and mass are in direct relationship*". Halm was soon followed in this concept by Ejnar Hertzsprung who, in 1919, also established a relationship between these two variables [94]. An excellent historical review on the subject exists [95]. For theoretical astrophysics, Eddington's confirmation of the mass-luminosity relationship was not simply an affirmation of Halm and Hertzsprung [93, 94]. It represented the birth of the fully gaseous Sun and of theoretical astrophysics.

The derivation of the mass-luminosity relationship would become a direct confirmation that Eddington's entire approach was correct. Stars, it seemed, must be gaseous. The argument was powerful. Still, it remained strangely dissociated from all physical observations of the Sun itself. In order to reproduce the mass-luminosity relationship, Eddington had only one requirement: the line he would draw would be guided by passing through a single star — Capella [92]. Jeans was not convinced. In 1925, he argued that the mass-luminosity relationship itself was nothing but an illusion: "*. . . there is no general relation between the masses and luminosities of stars. . .* " [85, p. 67].

Despite Jeans' objection, Eddington was quick to gain broad acceptance of his views. He would soon write a highly read popular work, *Stars and Atoms* [96]. It would provide a powerful look at both his philosophy and his scientific positions. In *Stars and Atoms*, Eddington stated that "*The Sun's material, in spite of being denser than water, really is a perfect gas. It sounds incredible, but it must be so*" [96, p. 38]. Further, Eddington would invoke Ralph Fowler in claiming that the gas was "*superperfect*" and "*more easily compressed than an ordinary gas*" [96, p. 40]. He would go on to state: "*It is now well realized that the stars are a very important adjunct to the physical laboratory — a sort of high temperature annex where the behavior of matter can be studied under greatly extended conditions. Being an astronomer, I naturally put the connexion somewhat differently and regard the physical laboratory as a low temperature station attached to the stars. In it the laboratory conditions which should be counted as abnormal*" [96, p. 83]. These words, of course, echoed Jeans' claim that Eddington had abandoned the laws of Earthly physics. Milne was forceful regarding Eddington: "*No words are needed to praise Eddington's achievement in calculating the state of equilibrium of a given mass of gas, and in calculating the rate of radiation from its surface. What was wrong was Eddington's failure to realize exactly his achievements: he had found a condition for a star to be gaseous throughout; by comparison with the star, Capella, he had evaluated the opacity in the boundary layers; and he had made it appears unlikely that the stars in nature were gaseous throughout. His claims were the contrary; he claimed to have calculated the luminosity of the existing stars; he claimed to show that they were gaseous throughout; and he claimed to have evaluated the internal opacity of the stars. Jeans deserves great credit for being the first critic to be skeptical about these claims of Eddington's theory, in spite of the attractive plausibility with which the theory was expounded*" [84, p. 27].

Recently, Alan Whiting presented a review of *Stars and Atoms* [97, p. 215–229]. Whiting claimed that Eddington was carefully aware of observational physics, particularly with re-

gards to the mass-luminosity question [97]. Whiting created an interesting contrast with Stanley [85] relative to the Jeans-Eddington battle. Whiting was highly critical of Jeans, but much more reverential towards Eddington [97, p. 215–229]. Perhaps this was with good reason as Eddington had championed the gaseous stars. This was to become the prevailing theory. Jeans defended the liquid alternative [97, p. 187–214]. Eventually though, even Jeans abandoned the liquid [97, p. 231–246] in favor of Eddington's gaseous models.

3.2 James Hopwood Jeans

Milne said of Jeans that "*he never wrote a dull page of mathematics in his life*" [84, p. 15]. Thus, in every respect, Jeans was a fitting adversary for Eddington. While an undergraduate at Cambridge, he received outstanding scores on his entrance exams to Trinity College and, along with G. H. Hardy, he would become the first student to take Part I of the Mathematical Tripos in only two years [84, p. 4–5]. A brilliant mathematician, Jeans' first great contribution to theoretical physics would be his study of the partition of energy between matter and radiation [98–100]. The papers demonstrated that Planck's quantum mechanical formulation [101], devoid of the Jeans-Rayleigh ultra-violet catastrophe, was the proper solution to the blackbody problem. Milne reviewed Jeans' contribution to the energy partition problem [84, p. 89–98]. Milne also provided perhaps the best condensed review of Jeans' position on liquid stars [84, p. 99–124]. In doing so, he reminded us that one of Jeans most beautiful works was his Adams Prize Essay [102]: "*Jeans Adams Prize Essay of 1919 was and remains a classic, even where subsequent discoveries have proved it wrong*" [84, p. 57]. The Essay was Jeans' first great venture into liquid stars.

Jeans was not the first to consider the problem of rotating homogeneous masses. As shall be seen in Section 3.3, the problem had been addressed by many of the finest minds in science. For Jeans, this included Poincaré [103] and George Darwin [104–108], the Cambridge physicist who had judged the *Adams Prize Essay* [84, p. 11]. Schwarzschild had also devoted time to this problem [109] and his approach remains important [110].

For Jeans, the starting point for liquid stars appears to have been the observation that a very large portion of these bodies existed as binary systems. The prevalence of binary stars would open the *Adams Prize Essay* [102, p. 2–4]. It would become a central part of *Astronomy and Cosmogony* [111, p. 20–23] and of his popular *The Universe Around Us*, both in its First Edition of 1933 [112, p. 38–53] and in the dramatically different Fourth Edition of 1944 [113, p. 37–51]. Relative to the formation of binaries, he wrote: "*In brief every rotating body conducts itself either as if it were purely liquid, or as if it were purely gaseous; there are no intermediate alternatives. Observational astronomy leaves no room for doubt that a great number of stars, possibly even all stars, follow the sequence shown in fig. 11. No other mechanism, so far as we know; is available for the formation of the numerous spectroscopic binary systems, in which two constituents describe small orbits about one another. In these stars, then, the central condensation of mass must be below the critical amount just mentioned; to this extent they behave like liquids rather than gases*" [112, p. 215]. Figure 11 represented the pear-shaped Darwin sequence of stellar evolution.

Three major problems preoccupied Jeans: 1) the purely rotational problem of a homogenous liquid, 2) tidal problem wherein a primary mass was affected by a secondary object, and 3) the formation of binary stars and maintenance of binary stars [84, p. 110]. For Jeans, the entire problem of the stars was one of physical stability. His work on liquids was surprisingly sparse of the radiative considerations which had characterized Eddington's entire approach to gaseous stars.

Jeans argued in Astronomy and Cosmogony that gaseous stars were inherently unrealistic [111, p. 64–104]: "*. . . we investigated the internal equilibrium of the stars on the supposition that they were masses of gravitating gas, in which the gas-laws were obeyed throughout. The investigation was abandoned when it was found to lead to impossibly high values of atomic weights of the stellar atoms. This created a suspicion that the hypothesis on which it was based was unfounded, and that the gas-laws are not obeyed in stellar interiors*" [111, p. 136]. He had previously attacked the stability of gaseous stars in the 1925 *Monthly Notices* [114]. He claimed that stars which generate energy as a function of temperature and density, would be violently unstable to radial oscillations [114]. Cowling refuted Jeans' claims [115, 116] and Whiting recently followed suit [117]. In the end, the instability of gaseous stars would survive scrutiny.

By the time Astronomy and Cosmogony was published, Jeans still refused to accept that the mass luminosity relationship was valid [111, p. 83]. Rather, he held that the mass-luminosity law could not be real, but that it was "*a consequence merely of the special assumption that kG is constant, and cannot have reference to actual stellar conditions*" [111, p. 83]. Jeans viewed the entire relation as a mathematical trick [85, p. 75]. Already, Jeans believed that stars were driven by the fission of materials such as uranium [111, p. 83]: "*But if the star has a liquid, or partially liquid, centre, this strip of safe land is so wide that, consistently with stability, the stellar material may have exactly the property that we should à priori expect to find, namely that its annihilation proceeds, like radio-active disintegration, at the same rate at all temperatures. If the substance of the star has this property, the star can no longer be in danger of exploding, for a mass of uranium or radium does not explode whatever we do to it*" [112, p. 287]. The amount of emitted light depended on the nature of the stellar constituents, not on a star's mass. Still, Jeans did not relate the ability to emit radiation to the phases of matter.

When Jeans first wrote *The Universe Around Us* [112], he postulated that, in order for a star to be stable, it must contain,

at the minimum, a liquid central region: "*And mathematical analysis shews that if the centre of the star is either liquid, or partially so, there is no danger of collapse; the liquid center provides so firm a basis for the star as to render collapse impossible*" [112, p. 287]. He advanced two postulates: "*1. That the annihilation of stellar matter proceeds spontaneously, not being affected by the temperature of the star. 2. That the central regions of stars are not in a purely gaseous state; their atoms, nuclei and electrons are so closely packed that they cannot move freely past one another, as in a gas, but rather jostle one another about like the molecules of a liquid*" [112, p. 287]. Jeans' concept of a liquid star was based not only on the stability of the resulting structures, but also on its constitutive materials and the need to provide the energy dissipated in the Sun's thermal radiation.

In his *Hindsight and Popular Astronomy*, Whiting [97] addressed at length the differences between Jeans' two Editions of his classic text *The Universe Around Us* [112, 113, p. 83]. These two editions were drastically at odds with one another. The first made the case for liquid stars, while the second advocated gaseous entities. Jeans completely removed any reference to liquid stars from the index of the 1944 edition [113]. The listing had many entries in the previous editions. Thus, it appears that a great transformation occurred for Jeans between 1933 and 1944. The evolution of Jeans' ideas were not recorded in the scientific literature. Jeans' last technical paper [84, p. 60] was entitled: "*Liquid Stars, a Correction*" [118]. It was published in 1928 at the same time as *Astronomy and Cosmogony* [111], but did not address liquid stars. Rather, it tackled Jeans' concerns relative to the instability of gaseous stars.

Why did Jeans abandon liquid stars? The answer will probably remain elusive. It was clear that Jeans had advocated that liquid stars were constituted of heavy elements which derived their energy from fission. As a result, when evidence gathered that hydrogen was the principle constituent of stars like the Sun [119–121], Jeans was left without a building block and without a means to generate energy. It was inconceivable to a person in Jeans' day that hydrogen could exist in liquid form, provide the requisite building material for a liquid star, and maintain the Sun's energy through fusion [56]. Furthermore, Jeans had to contend with the critical temperature arguments based on Andrews [29]. Given the need for hydrogen, it must have seemed to Jeans that liquid stars were doomed.

3.3 Subrahmanyan Chandrashekhar and rotating fluid masses

Subrahmanyan Chandrasekhar (October 13th, 1910 — August 21, 1995) [122] was Ralph Fowler's student at Cambridge. He was well acquainted with Eddington, Jeans, and Milne. Eventually, he would become the recipient of the 1983 Nobel Prize in physics. His text, *Introduction to the Study of Stellar Structure* remains an authoritative treatment of the subject matter and is widely considered a classic in astrophysics [48]. Chandrasekhar also wrote a lesser known volume on *Ellipsoidal Figures of Equilibrium* [124]. Rotating fluid masses captivated Chandrasekhar for a period of nine years [124, p. 241]. The father of modern solar astrophysics makes two points with regards to his time investment: 1) "*the subject had attracted the attention of a long succession of distinguished mathematicians and astronomers*" and 2) "*the method of the virial is not restricted to homogeneous masses*" [124, 241].

Except for a single chapter, *Ellipsoidal Figures of Equilibrium* was entirely devoted to homogeneous liquid masses. His *Historical Introduction* [124, p. 241] provided a magnificent review of the field which outlined the seminal contributions of men like Newton, Maclaurin, Jacobi, Meyer, Liouville, Dirichlet, Dedekind, Riemann, Poincaré, Cartan, Roche, Darwin, and Jeans.

Chandrasekhar believed that the problem of the homogeneous liquid mass "*had been left in an incomplete state with many gaps and omissions and some plain errors and misconceptions*" [124, p. 241]. This was the prime motivation for his text. The most significant gap in the theory of the homogeneous rotating liquid was addressed with Chandrasekhar's discussion of the Darwin ellipsoids [124, p. 218–239]. In a chapter devoted to the Roche ellipsoids, he demonstrated that such structures are unstable over the entire Darwin sequence [124, p. 218–239]. Chandrasekhar's conclusion was a partial setback for Jeans' work, in that the latter had speculated, as seen in Section 3.2, that binaries were formed through the evolution of the Darwin sequence [112, p. 247–253]. Both Jeans and Darwin had recognized that the pear-shaped figure was unstable [112, p. 252], though they did not suspect that this was the case for the entire sequence. As a result, the extensive presence of binaries in the sky, Jeans' primary argument for liquid stars, could not be easily explained by the liquid models he had advocated after all. Relative to binaries, it seems that neither liquid nor gaseous models have offered a definitive answer. Lebovitz argued that "*the viability of fission theory remains unsettled to this day*" [125, p. 131].

4 Conclusions

Throughout the ages, as new physical discoveries occurred, attempts were made to mold them into the prevailing model of our star. Secchi's Sun, with its particulate photospheric matter floating on a gaseous globe, was not easily abandoned [38, 39]. Faye's insistence that the Sun was devoid of a true surface has remained accepted to this day [2]. Stoney's sprinkling of graphite particles on the Sun would prevail for 60 years [11]. But when Stoney was eventually abandoned, could modern man really endow a gas with features found only in condensed matter? Could the solar spectrum truly be accounted for by the mathematics linked to gaseous stars? These were the questions that begged for answers, although

they could not be resolved solely through historical review. They would require instead a careful analysis of the stellar opacity problem [80].

It has always been true that current solar models far surpass in validity those advanced by previous generations. Therefore, modern science must be called to greater caution. It is noteworthy that, while Laplace's nebular hypothesis and Helmholtz' contraction theory have long ago been abandoned [1], the influence they carried in forging a gaseous Sun did not wane. In like manner, Kirchhoff's law of thermal emission [9, 73], though never validated in a gas, has remained a pillar of modern solar theory [1]. This has been the case, even though no gas has ever emitted a continuous spectrum which varied as the 4th power of temperature. Thermal emissivities in gases tended to drop with temperature, not to dramatically increase [89]. Invoked as one of the early pillars of the gaseous Sun, the broadening of hydrogen has never assumed a blackbody line shape. In the gaseous state, despite increased pressure, hydrogen cannot emit with a 4th power relationship [89]. In 1869, Andrews [29] was unaware that liquid metallic hydrogen existed [56]. The existence of this material [56], has delivered a devastating defeat to the limiting aspect of critical temperatures [29] measured in ordinary gases, relative to forming a gaseous Sun [1]. Given these considerations, what can be said about our solar models?

With the publication of Arthur Eddington's *Internal Constitution of the Stars* [3] and the subsequent work *An Introduction to the Study of Stellar Structure* by Subrahmanyan Chandrasekhar [48], astrophysics seemed to have taken unprecedented steps in understanding the stars. Eddington's classic work advanced a cohesive gaseous model. It also brought forth the phenomenal mass-luminosity relation, so prized by theoretical astrophysics. For his part, Chandrasekhar would propel our knowledge of stellar evolution with his introduction of degeneracy and his tremendous treatment of the white dwarf, leading to the limit which bears his name [48]. Given the powerful theoretical framework which surrounded the gaseous stars, most envision that a perfect marriage of physical observation and mathematical prowess had resulted in a level of sophistication well beyond that reached in ages past.

In spite of all this, as a celestial body, the Sun has structure: a photosphere, a chromosphere, a corona, granulations, sunspots, prominences, etc. However, by their very nature, gases are unable to impart structure. Long ago, Jeans complained that "*All of Eddington's theoretical investigations have been based on assumptions which are outside the laws of physics*" [88]. The criticism may be overly harsh, but it must be remembered that many astronomers of the period, unlike Eddington, placed a strong emphasis on physical observation. For his part, Eddington essentially dismissed physical findings. Hence, it is not surprising that animosity arose between these two men. As the author previously stated: "*Eddington believed that the laws of physics and thermodynamics could be used to deduce the internal structure of the Sun without any experimental verification. In 1926, he would speak hypothetically about being able to live on an isolated planet completely surrounded by clouds. In such a setting, he thought he would still be able to analyze the Sun without any further knowledge than its mass, its size and the laws of physics*" [126]. Eddington himself realized the risks he was taking when he wrote that: "*We should be unwise to trust scientific inferences very far when it becomes divorced from opportunity for observational tests*" [44, p. 1]. Since Eddington was trying to understand stellar interiors, there could be no observational confirmation of his mathematics. In addition, Eddington's treatment completely sidestepped the structural features on the Sun. Moreover, Eddington assumed the same average coefficient of absorption throughout a star despite fluctuations in temperatures and densities [44]. He treated all opacities, for both dense stars and sparse ones, as corresponding to the opacity within the Sun itself [44]. His model could not be tested using data from the Sun.

Eddington sought to establish the mass-luminosity relationship as a manifestation that at least some merit could be gained from his approach. This relationship was enticing, but its acceptance would come at a great price. Theoretical astrophysics would be brought to the uncomfortable position of minimizing the importance of direct physical evidence for the state manifested by the Sun. This was the cost of embracing stellar, rather than solar, data. Direct solar observations received less weight than distant stellar findings. This was the case even though stellar measurements were obtained, following assumptions and manipulation from stars positioned light years, if not thousands of light years, away. Additionally, by adopting Eddington's conclusion, the chemical nature of the star itself was quietly dismissed as immaterial [44]. Yet on Earth, the thermal emission of all materials was determined strictly by their chemical makeup and physical structure [127]. These facts should not be overlooked. It was improper for Eddington to discount earthly laboratories, as seen in Section 3.1, because mankind could trust no other venue.

If Eddington struggled in certain areas, his approach was not without precedent. As described earlier [1], those who studied solar physics, from Galileo to Wilson to Herschel to Spencer to Secchi and Faye, had no alternative course of action. Eddington was correct: given our limitations, educated speculation was the only avenue. Furthermore, it would prove much easier, in making progress in science, to rebuke known ideas, rather than to speculate on the unknown. Eddington's attempt to forge new ground was laudable and such will remain the case through the ages.

Though Jeans philosophically disagreed with Eddington's approach [85], he was unable to truly offer an alternative. Many of his claims were incorrect. He continued to believe in Helmholtz' theory of contraction for energy production, well after many had abandoned the idea [85]. He advocated liquid stars as a mechanism for producing binaries, when more prudent mathematical treatments would cast doubt upon his argu-

ments [124]. He advocated that gaseous stars were unstable to oscillations [114]. He advanced that liquid stars had to be formed from uranium and radium [112, p. 287]. In the battle with Eddington, he showed a lack of restraint in charging that his colleague's approaches were not even science. Who, from sole authority, could establish what was or was not science? Rather, as Milne highlighted, Jeans and Eddington should have made a concerted effort to work together [84, p. 28]. The questions were much too complex for isolated approaches and both men would have been well served to collaborate.

As this review of the *Thermodynamic History of the Solar Constitution* comes to a close, one can only wonder at the beauty of solar science. Stellar astrophysics remains a relatively small island in the sea of science. Nonetheless, so many aspects of earthly physics and chemistry touch the subject. In this regard, and given the task ahead, there is much to contribute to the subject area, even for non-astronomers. Thus, we leave the subject by pondering, once again [1], upon the wisdom offered by the magnificent solar astronomer, George Hale [128]. In writing the obituary for Arthur Schuster [24], the founder of the *Astrophysical Journal* [128] was sickly and approaching the end of his own life. Hale reminded us of the need to work together in order to arrive at a deeper understanding of the world around us. A study of the history of solar science echoes Hale. The contributions of many were required to arrive at some semblance of the truth: "*A Galileo or a Newton or an Einstein cannot be produced by an International conference, nor can lesser men who have nevertheless contributed enormously to original thought. How then are we to reconcile our co-operative projects with the prime necessity for personal freedom?* [24, p. 101] ... "*One of the most important needs of science is to establish closer relationships between workers in different fields. It is comparatively easy to bring together specialists in given subjects and to secure their friendly co-operation. But to fill the gaps between various branches of science is a more difficult task, in spite of the obvious possibilities of advance. Such possibilities are shown by the development of astrophysics, geophysics, biochemistry, and many other subjects. However, the fact remains that countless opportunities are lost because instruments, methods, and ideas which have originated in some particular field are unknown or at least unused in other fields*" [24, p. 102].

Notes and acknowledgements

The author would like to thank Bernadette Carstensen for the graphite illustrations of James Hopwood Jeans and Arthur Stanley Eddington. These illustrations are used with permission. All translations from French were generated by the author. Dmitri Rabounski is recognized for first highlighting to the author the important contributions of James Jeans.

Dedication

This work is dedicated to my youngest son, Luc.

Submitted on May 13, 2011 / Accepted on May 18, 2011
First published online on May 31, 2011 / Revised on June 12, 2011

References

1. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v.3, 3–25.
2. Faye H. Sur la constitution physique du Soleil. *Les Mondes*, 1865, v.7, 293–306 (translated into English by Patrice Robitaille: On the Physical Constitution of the Sun — Part I, *Progr. Phys.*, 2011, v.3, 35–40 — a paper published in this Special Issue).
3. Eddington A.S. The internal constitution of the stars. *The Observatory*, 1920, v.43 (557), 341–357.
4. Glockler G. The heat of sublimation of graphite and the composition of carbon vapor. *J. Chem. Phys.*, 1954, v.22, 159–161.
5. Wedgwood T. Experiments and observations on the production of light from different bodies, by heat and by attrition. *Phil. Trans. Roy. Soc. Lond.*, 1792, v.82, 28–47; Wedgwood T. Continuation of a paper on the production of light and heat from different bodies. *Phil. Trans. Roy. Soc. Lond.*, 1792, v.82, 270–282.
6. Leslie J. An Experimental Inquiry into the Nature and Propagation of Heat. T. Gillet, Salisbury Square, 1804.
7. Robitaille P.M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v.3, 36–55.
8. 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(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).
9. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsverögen und dem Absorptionsvermögen. der Körper für 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.*, ser. 4, 1860, v.20, 1–21).
10. Stewart B. Lessons in Elementary Physics (New Edition). MacMillan and Co., London, 1884.
11. Stoney G.J. On the physical constitution of the sun and stars. *Proc. Roy. Soc. London*, 1867, v.16, 25–34
12. Huggins W. Inaugural Address. *Nature*, 1891, v.44(1138), 372–382.
13. Maxwell J.C. The Theory of Heat (3rd Ed.). Longmans, Green, and Co., London, 1872.
14. Stewart B. Lessons in Elementary Physics. MacMillan and Co., London, 1884.
15. de Groot J. J., van Vliet J. A. J. M. The High-Pressure Sodium Lamp. Kluwer Technische Boeken BV, Deventer, 1986.
16. Wilson W.E. On the effect of pressure of the surrounding gas on the temperature of the crater of an electric arc light. Preliminary notes of observations made at Daramona, Streete, co. Westmeath. *Proc. Roy. Soc. London*, 1895, v.58, 174–176; also in *Nature*, 1895, v.52(1340), 238; also in *Astrophys. J.*, 1895, v.2, 212–214.
17. Wilson W.E. and Fitzgerald G.F. On the effect of pressure in the surrounding gas on the temperature of the crater of an electric arc. Correction of results in former paper. *Astrophys. J.*, 1897, v.5, 101–108.
18. Ball R. The sun's light. *McClure's Magazine*, 1896, v.6(2), 106–110.
19. Newcomb S. Popular Astronomy. Harper & Brothers, New York, 1884.
20. Young C.A. A Textbook of General Astronomy for Colleges and Scientific Schools. Ginn & Co., Boston, 1909.

21. Laudauer J. and Tingle J.B. Spectrum Analysis. John Wiley & Sons, New York, 1907.
22. Maunder E.W. Are the Planets Inhabited? Harper & Brothers, London, 1913.
23. Roscoe H.E. and Schuster A. Spectrum Analysis — Six Lectures Delivered in 1868 before the Society of Apothecaries of London. 4th Edition, Macmillan and Co., London, 1885.
24. Hale G.E. Sir Arthur Schuster. *Astrophys. J.*, 1935, v.81(2), 97–106.
25. Unsigned. Sir Arthur Schuster, F.R.S. *Mon. Not. Roy. Astro. Soc.*, 1935, v.95, 326–329.
26. Schuster A. Radiation through a foggy atmosphere. *Astrophys. J.*, 1905, v.21, 1–22.
27. Schuster A. Radiation through a foggy atmosphere [abridged]. *The Observatory*, 1903, v.26, 379–381.
28. Clapeyron E. Mémoire sur la puissance motrice de la chaleur. *Journal de l'École Polytechnique*, 1834, v.14, 153–190.
29. Andrews T. The Bakerian lecture: On the continuity of the gaseous and liquid states of matter. *Phil. Trans. Roy. Soc.*, 1869, v.159, 575–590.
30. Fison A.H. Recent Advances in Astronomy, Blackie and Son, London, 1898.
31. Zöllner F. On the temperature and physical constitution of the Sun. *Phil. Mag.*, 4th series, 1870, v.40, 313–327 (essentially reprinted in: Zöllner F. On the Sun's temperature and physical constitution. *Nature*, 1870, v.2(52), 522–526).
32. Zöllner F. On the temperature and physical constitution of the Sun (Second Memoire). *Phil. Mag.*, 4th series, 1873, v.46(306), 290–304.
33. Scheiner J. The temperature of the Sun — I. *Astron. Soc. Pac.*, 1898, v.9(64), 167–179.
34. Scheiner J. The temperature of the Sun — II. *Astron. Soc. Pac.*, 1898, v.10(65), 224–234.
35. Unsigned (Spencer H.) Recent astronomy, and the nebular hypothesis. *Westminster Review*, 1858, v.70, 185–225.
36. Spencer H. The constitution of the sun. *The Reader: A Review of Current Literature*, 25 February 1865, v.5, 227–229 (also found in: *Spencer H. Essays: Scientific, Political, and Speculative*. Vol. III. Williams and Norgate, London, 1875, 217–229).
37. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H. E. Roscoe, Macmillan and Co., Cambridge, 1862, 23–31.
38. Secchi A. Sulla Teoria Delle Macchie Solari: Proposta dal sig. Kirchoff. *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 31 January 1864, v.3(4), 1–4 (translated into English by Eileen Reeves and Mary Posani: On the Theory of Sunspots Proposed by Signor Kirchoff, *Progr. Phys.*, 2011, v.3, 26–29 — a paper published in this Special Issue).
39. Secchi A. Sulla Struttura della Fotosfera Solare. *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 30 November 1864, v.3(11), 1–3. (translated into French by François Moigno: Sur la structure de la photosphère du soleil. *Les Mondes*, 22 December 1864, v.6, 703–707; translated into English by Mary Posani and Eileen Reeves: On the structure of the photosphere of the sun. *Progr. Phys.*, 2011, v.3, 30–32 — a paper published in this Special Issue).
40. Faye H. Sur la constitution physique du Soleil — Première Partie. *Les Mondes*, 1865, v.7, 293–306 (translated into English by Patrice Robitaille: On the Physical Constitution of the Sun — Part I. *Progr. Phys.*, 2011, v.3, 35–40 — a paper published in this Special Issue).
41. Abbe C. Memoir of Jonathan Homer Lane — Read before the Academy in April 1892. *Biographical Memoirs, National Academy of Sciences*, 1895, v.iii, 253–264.
42. Lane J.H. On the theoretical temperature of the Sun; under the hypothesis of a gaseous mass maintaining its volume by its internal heat, and depending on the laws of gases as known to terrestrial experiments. *Am. J. Science*, 2nd Ser. 1 (1870), v.50, 57–74 (Also found in Meadows A.J. Early solar physics, New York, 1970, 254–276).
43. Stevenson-Powell C. J. Homer Lane and the Internal Structure of the Sun. *J. Hist. Astron.*, 1988, v.19, 183–199.
44. Eddington A.S. The Internal Constitution of the Stars. Dover Publ. Inc., New York, 1959.
45. Ritter A. Untersuchungen über die Höhe der Atmosphäre und die Constitution gasförmiger Weltkörper. *Wiedemann's Annalen der Physik und Chimie*, 1878, v.6, 135 (Most cited of a series of papers in the same journal — citations for the others, taken from [48], are as follows: 1878, v.5, 405; 1878, v.5, 543; 1879, v.7, 304; 1880, v.8, 157; 1880, v.10, 130; 1880, v.11, 332; 1880, v.11, 978; 1881, v.12, 445; 1881, v.13, 360; 1881, v.14, 610; 1882, v.16, 166; 1882, v.17, 332; 1883, v.18, 488; 1883, v.20, 137; 1883, v.20, 897; 1883, v.20, 910).
46. Thomson W. On the equilibrium of a gas under its own gravitation only. *Phil. Mag.*, 1887, v.23, 287–292
47. Emden R. Gaskugeln: Andwendungen der Mechanischen Wärmetheorie auf Kosmologische und Meteorologische Probleme (Gas spheres: Applications of the mechanical heat theory to cosmological and meteorological problems), Druck und Verlag von B.G. Teubner, Leipzig, 1907.
48. Chandrasekhar S. Introduction to the Study of Stellar Structure. Dover Publications, (University of Chicago Press reprint), 1957.
49. Clausius R. Über die bewegende Kraft der Wärme — Part I and Part II. *Poggendorfs Annalen der Physik und Chemie*, 1850, v.79, 368–397 and 500–524 (also in English: On the moving force of heat, and the laws regarding the nature of heat itself which are deducible therefrom. *Phil. Mag.*, 1851, v.2, 1–21 and 102–119).
50. Clausius R. The Mechanical Theory of Heat — with its Applications to the Steam Engine and to Physical Properties of Bodies. John van Voorst, London. 1865.
51. Thomson W. On the dynamical theory of heat; with numerical results deduced from Mr. Joule's equivalent of a thermal unit and M. Regnault's observations on steam. *Trans. Roy. Soc. Edinburgh*, 1851, v.20, 261–288.
52. Peirce B. Ideality in the Physical Sciences. Boston, 1881.
53. See T.F. F. On the heat of the Sun and on the temperatures and relative ages of the stars and nebulae. *Astronomische Nachrichten*, 1898, v.148, 179–184.
54. See T.F.F. The fundamental law of temperature for gaseous celestial bodies. *Astronomical J.*, 1899, v.19(23), 181–185.
55. Hastings C.S. A theory of the constitution of the Sun, founded upon spectroscopic observations, original and other. *Am. J. Science*, 1881, v.21(121), 33–44.
56. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v.3 — paper published in this Special Issue.
57. Very F.W. The absorptive power of the solar atmosphere. *Astrophys. J.*, 1902, v.16, 73–91.
58. Ogden J.G. and Very F.W. *Popular Astronomy*, 1928, v.36(7), 389–397.
59. Langley S.P. Experimental determination of wave-lengths in the invisible spectrum. *Mem. Natl. Acad. Sci.*, 1883, v.2, 147–162.
60. Langley S.P. On hitherto unrecognized wave-lengths. *Phil. Mag.*, 1886, v.22, 149–173.
61. Langley S.P. The invisible solar and lunar spectrum. *Am. J. Science*, 1888, v.36(216), 397–410.
62. Schuster A. The solar atmosphere. *Astrophys. J.*, 1902, v.16, 320–327.

63. Meadows A.J. Early Solar Physics. Pergamon Press, Oxford, 1970.
64. Very F.W. The absorption of radiation by the solar atmosphere and the intrinsic radiation of that atmosphere. *Astrophys. J.*, 1904, v.19, 139–150.
65. Schuster A. The temperature of the solar atmosphere. *Astrophys. J.*, 1905, v.21, 258–260.
66. Menzel D.H. Selected Papers on the Transfer of Radiation: 6 papers by Arthur Schuster, K. Schwarzschild, A.S. Eddington, S. Rosseland, E.A. Milne. Dover Publications, Inc., New York, 1966.
67. Chandrasekhar, S. Radiative Transfer. Clarendon Press, Oxford, 1950.
68. Rybicki G.B. and Lightman A.P. Radiative Processes in Astrophysics, Wiley-VCH, Weinheim, Germany, 2004.
69. Swihart T.L. Radiation Transfer and Stellar Atmospheres (Astronomy and astrophysics series), Tucson : Pachart, 1981.
70. Milne E.A. Bakerian Lecture: The structure and opacity of a stellar atmosphere. *Phil. Trans. Roy. Soc. London*, 1929, v.228, 421–461.
71. Watts W.M., Huggins W. An Introduction to the Study of Spectrum Analysis. Longmans, Green, and Co., London, 1904.
72. Robitaille P.M. A critical analysis of universality and Kirchhoff's law: a return to Stewart's law of thermal emission. *Progr. Phys.*, 2008, v.3, 30–35; arXiv: 0805.1625.
73. Robitaille P.M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v.4, 3–13.
74. Sampson R.A. On the rotation and mechanical state of the Sun. *Memoires Roy. Astron. Soc.*, 1894, v.51, 123–183.
75. Unsigned. Karl Schwarzschild. *The Observatory*, 1916, v.39(503), 336–339.
76. Schwarzschild K. Über das Gleichgewicht der Sonnenamospäre. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttigen, Göttinger Nachrichten*, 1906, 195, 41–53 (English translation: Concerning the equilibrium of the solar atmosphere, found in: Menzel D.H. [66], p.25–34 and Meadows A.J. [13], 277–290).
77. Eddington A.S. On the radiative equilibrium of the stars. *Mon. Not. Roy. Astron. Soc.*, 1916, v.77(1), 16–35 (Also found in Lang K.R. and Gingerich O.: A Source Book in Astronomy and Astrophysics, 1900–1975. Harvard University Press, Cambridge, MA, 1979, p.225–235; also found in [66], p. 53–72).
78. Rosseland S. Note on the absorption of radiation within a star. *Mon. Not. Roy. Astron. Soc.*, v.84(7), 525–528.
79. Rosseland S. Theoretical Astrophysics: Atomic Theory and the Analysis of Stellar Atmospheres and Envelopes. Clarendon Press, Oxford, 1939.
80. Robitaille P.M. Stellar opacities: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v.3 — a paper in this Special Issue.
81. Chandrasekhar S. Eddington: The Most Distinguished Astrophysicist of his time. Cambridge University Press, 1983.
82. McCrea W.H. Edward Arthur Milne. *The Observatory*, 1950, v.70, 225–232.
83. Chandrasekhar S. Ralph Howard Fowler. *Astrophys. J.*, 1945, v.101(1), 1–5.
84. Milne E.A. Sir James Jeans — A Biography. Cambridge University Press, Cambridge, 1952.
85. Stanley M. So simple a thing as a star: The Eddington-Jeans debate over astrophysical phenomenology. *Brit. J. Hist. Science*, 2007, v.40(1), 53–82.
86. Rosenthal-Scheider I. Reality and Scientific Truth. Wayne State University Press, Detroit, 1949.
87. Dyson F.W., Eddington A.S. and Davidson C.R. A determination of the deflection of light by the Sun's gravitational field, from observations made at the solar eclipse of May 29, 1919. *Phil. Trans. Roy. Soc.*, 1920, v.220, 291–333.
88. Unsigned. Meeting of the Royal Astronomical Society — Friday, 1925 January 9. *The Observatory*, 1925, v.48, 29–38.
89. Robitaille P.M. The little heat engine: Heat transfer in solids, liquids, and gases. *Progr. Phys.*, 2007, v.4, 25–33.
90. Stefan J. Stefan J. Ueber die Beziehung zwischen der Warmestrahlung und der Temperatur. *Wein. Akad. Sitzber.*, 1879, v.79, 391–428.
91. Jeans J.H. The equations of radiative transfer of energy. *Mon. Not. Roy. Astron. Soc.*, 1917, v.78, 28–36.
92. Eddington A.S. On the relation between the masses and the luminosities of the stars. *Mon. Not. Roy. Astron. Soc.*, 1924, v.84, 308–322 (Also found in Lang K.R. and Gingerich O.: A Source Book in Astronomy and Astrophysics, 1900–1975. Harvard University Press, Cambridge, MA, 1979, p.291–302).
93. Halm J. Further considerations relating to the systematic motions of the stars. *Mon. Not. Roy. Astron. Soc.*, 1911, v.71, 610–639.
94. Hertzsprung E. Bemerkungen zur Statistik der Stern parallaxen. *Astronomische Nachrichten*, 1919, v.208, 89–96; also Bemerkungen zur Spektralklassifikation der gelben Stern. *Astronomische Nachrichten*, 1919, v.208, 265–272.
95. Nielsen A.V. Contributions to the history of the Hertzsprung-Russell diagram. *Centaurus*, 1964, v.9(4), 219–253.
96. Eddington A.S. Stars and Atoms. Yale University Press, New Haven, 1927.
97. Whiting A.B. Hindsight in Popular Astronomy. World Scientific, New Jersey, 2011.
98. Jeans J.H. On the partition of energy between matter and aether. *Phil. Mag.*, Series 6, 1905, v.10(55), 91–98.
99. Jeans J.H. A comparison between two theories of radiation. *Nature*, 1905, v.72, 293–294.
100. Jeans J.H. On the laws of radiation. *Proc. Roy. Soc. London A*, 1905, v.76(513), 545–552.
101. Planck M. Ueber das Gesetz der Energieverteilung in Normalspectrum. *Annalen der Physik*, 1901, v.4, 553–563.
102. Jeans J.H. Problems of Cosmogony and Stellar Dynamics — Being an Essay to which the Adams Prize of the University of Cambridge for the year 1917 was Adjudged. Cambridge University Press, 1919.
103. Poincaré H. Figures d'équilibre d'une masse fluide. Gauthier-Villars et Cie, Paris, 1902.
104. Darwin G.H. On Jacobi's figure of equilibrium for a rotating mass of fluid. *Proc. Royal Soc. London*, 1886, v.41, 319–336.
105. Darwin G.H. On figures of equilibrium of rotating masses of fluid. *Phil. Trans. Roy. Soc. London A*, 1887, v.178, 379–428.
106. Darwin G.H. The stability of the pear-shaped figure of equilibrium of a rotating mass of liquid. *Phil. Trans. Roy. Soc. London A*, 1903, v.200, 251–314.
107. Darwin G.H. On the figure and stability of a liquid satellite. *Phil. Trans. Roy. Soc. London A*, 1906, v.206, 161–248.
108. Darwin G.H. Further consideration of the stability of the pear-shaped figure of a rotating mass of liquid. *Phil. Trans. Roy. Soc. London A*, 1908, v.208, 1–19.
109. Schwarzschild K. On the gravitational field of a sphere of incompressible fluid according to Einstein's theory. *Sitzungsber. Preuss. Akad. Wiss., Phys. Math. Kl.*, 1916, 424–434 (English translation by L. Borissova and D. Rabounski: *The Abraham Zelmanov Journal*, 2008, v.1).
110. Crothers S. On the vacuum field of a sphere of incompressible fluid. *Progr. Phys.*, 2005, v.2, 76–81.
111. Jeans J.H. Astronomy and Cosmogony. Cambridge University Press, 1928.
112. Jeans J.H. The Universe Around Us. 1st Edition, Cambridge University Press, 1933.

113. Jeans J.H. The Universe Around Us. 4th Edition, Cambridge University Press, 1946.

114. Jeans J.H. A theory of stellar evolution. *Mon. Not. Roy. Astron. Soc.*, 1925, v.85, 914–933.

115. Cowling T.G. The stability of gaseous stars. *Mon. Not. Roy. Astron. Soc.*, 1934, v.94, 768–782.

116. Cowling T.G. The stability of gaseous stars — Second Paper. *Mon. Not. Roy. Astron. Soc.*, 1935, v.96, 42–60.

117. Whiting A.B. Sir James Jeans and the stability of gaseous stars. *The Observatory*, 2007, v.127(1), 13–21.

118. Jeans J.H. Liquid stars, a correction. *Mon. Not. Roy. Astron. Soc.*, 1928, v.88, 393–395.

119. Payne C.H. The relative abundances of the elements. In: *Stellar Atmospheres*, Harvard University Monograph no. 1, Harvard University Press, 1925, chapter 13 (Also found in Lang K.R. and Gingerich O.: A Source Book in Astronomy and Astrophysics, 1900–1975. Harvard University Press, Cambridge, MA, 1979, p.244–248).

120. Unsöld A. Über die Struktur der Fraunhofersehen Linien und die quantitative Spektralanalyse der Sonnenatmosphäre. *Zeitschrift für Physik*, 1928, v.46, 765–781.

121. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v.70, 11–82.

122. Garstang R.H. Subrahmanyan Chandrasekhar (1910–1995). *Publications of the Astronomical Society of the Pacific*, 1997, v.109, 73–77.

123. McCrea W. Obituary: Subramanyan Chandrasekhar. *The Observatory*, 1996, v.116, 121–124.

124. Chandrasekhar S. Ellipsoidal Figures of Equilibrium. Yale University Press, New Haven, 1969.

125. Srinivasan G. From White Dwarfs to Black Holes: The Legacy of S. Chandrasekhar. The University of Chicago Press, Chicago, 1999.

126. Robitaille P.M.L. The collapse of the Big Bang and the gaseous Sun. *New York Times*, March 17, 2002, page A10 (available online: http://thermalphysics.org/pdf/times.pdf).

127. Touloukian Y.S. and Ho C.Y. Thermophysical Properties of Matter (vols. 1–8). Plenum, New York, 1970.

128. Adams W.S. George Ellery Hale. *Astrophys. J.*, 1938, v.87(4), 369–388.

Liquid Metallic Hydrogen: A Building Block for the Liquid Sun

Pierre-Marie Robitaille
Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA
E-mail: robitaille.1@osu.edu

Liquid metallic hydrogen provides a compelling material for constructing a condensed matter model of the Sun and the photosphere. Like diamond, metallic hydrogen might have the potential to be a metastable substance requiring high pressures for formation. Once created, it would remain stable even at lower pressures. The metallic form of hydrogen was initially conceived in 1935 by Eugene Wigner and Hillard B. Huntington who indirectly anticipated its elevated critical temperature for liquefaction (Wigner E. and Huntington H. B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v.3, 764–770). At that time, solid metallic hydrogen was hypothesized to exist as a body centered cubic, although a more energetically accessible layered graphite-like lattice was also envisioned. Relative to solar emission, this structural resemblance between graphite and layered metallic hydrogen should not be easily dismissed. In the laboratory, metallic hydrogen remains an elusive material. However, given the extensive observational evidence for a condensed Sun composed primarily of hydrogen, it is appropriate to consider metallic hydrogen as a solar building block. It is anticipated that solar liquid metallic hydrogen should possess at least some layered order. Since layered liquid metallic hydrogen would be essentially incompressible, its invocation as a solar constituent brings into question much of current stellar physics. The central proof of a liquid state remains the thermal spectrum of the Sun itself. Its proper understanding brings together all the great forces which shaped modern physics. Although other proofs exist for a liquid photosphere, our focus remains solidly on the generation of this light.

1 Introduction

Decidedly, the greatest single impetus for a fully gaseous Sun [1, 2] was the elucidation of critical temperatures by Thomas Andrews in 1869 [3, 4]. Since ordinary gases could not be liquefied at the temperatures associated with the Sun, it was inconceivable that the photosphere was made from condensed matter: "*It is, however, scarcely possible to regard as existing in the interior of the Sun, matter in either the solid or in the liquid condition... Since, however, it became apparent from the classic research of Dr. Andrews in 1869, that there exists for every element a critical temperature, above which it is impossible for it under any conditions of pressure to assume the liquid state, it has generally been regarded that a liquid interior to the Sun is next to an impossibility*" [5, p. 36-37]. As a result of such logic, the idea that the Sun was gaseous flourished. Though Father Angello Secchi and Hervé Faye had already proposed a gaseous solar model [1], Andrews' discovery served to significantly validate their conjectures. Given the logic of the period, the body and photosphere of the Sun could not be liquid [1].

At the same time, scientists of the late 19th and early 20th century remained puzzled with respect to the solar spectrum [1, 2]. Because graphite was the prime source of blackbody radiation on Earth [6], G. Johnstone Stoney placed liquid or solid carbon on the surface of the Sun in 1867 [7]. It would remain there for the next 50 years [2]. Armed with graphite, it became simple to explain why the solar photosphere emitted a thermal spectrum resembling a blackbody. Over time, the enthusiasm for carbon began to wane. Charles Hastings argued that condensed carbon could not be present on the Sun. The temperatures involved did not permit such a hypothesis. Hastings required an alternative: "*At any rate, we are sure that the substance in question, so far as we know it, has properties similar to those of the carbon group*" [8]. Hastings did not elaborate on these properties, but it was clear that he was searching for a substance with unbelievable refractory characteristics, something with the structure of graphite. A material capable of producing the thermal spectrum of the Sun had to exist in the condensed state at tremendous temperatures.

Eventually, theoretical astrophysics dispensed of the need for condensed matter. In so doing, the stellar opacity problem was created [9]. It was Schuster's *Radiation through a Foggy Atmosphere* [10] which began to cast condensed matter out of the photosphere [2]. Schuster postulated that all gases, if sufficiently thick, emitted as blackbodies: "*The radiation in this case becomes equal to that of a completely black surface, which agrees with the well-known law that absorption irrespective of scattering tends to make the radiation of all bodies equal to that of a black body when the thickness is increased*" [10, p. 6]. Schuster's conclusion was not supported by the gaseous nebula. These celestial objects had long been known to emit line spectra [11, p. 87] and, though they were assuredly thick, blackbody lineshapes were not produced. As

previously outlined by the author [2], Schuster's error consisted in resting his derivation upon the premise that Kirchhoff's law of thermal emission was valid [12].

Gustav Kirchhoff insisted that, given thermal equilibrium with an enclosure, a blackbody spectrum could be produced by any object [12]. Yet, if Kirchhoff's law was correct, his contemporaries should not have refused to adopt a fully gaseous Sun throughout the 19th century [1, 2]. They would not have insisted on the need for graphite. If graphite was viewed as less than optimal, they would not have invoked pressure broadening as a means to produce the solar spectrum [1]. Kirchhoff's formulation, after all, was independent of pressure. It would become evident that something was not quite right with Kirchhoff's deductions. The author has outline why Kirchhoff's law of thermal emission was erroneous [13, 14]. On the simplest level, it constituted a violation of the first law of thermodynamics. In addition, as was outlined relative to the stellar opacity problem, gases remain unable to emit a blackbody spectrum [9]. This was the surest evidence that Kirchhoff's law was invalid.

As a result, if gases could not produce the solar spectrum, astrophysics should have returned to the condensed state. At the beginning of the 20th century, Jeans promoted liquid stars [15] based on stability arguments, only to discard them at the end of his life [2]. If Jeans abandoned liquids, it was likely due to his lack of a proper building block [2]. He conceived of stars as composed of heavy elements such as uranium and radium [2]. When the Sun was shown to contain large amounts of hydrogen [16–18], Jeans was left without a proper structural material. He did not anticipate that metallic hydrogen could exist [19] and that the substance provided the perfect candidate for a fully condensed Sun. In proposing the existence of metallic hydrogen [19], condensed matter physics would unknowingly provide Jeans with a suitable material for liquid stars [2]. Andrews' critical temperature in ordinary gases became inconsequential [20]. More intriguing was the observation that the layered lattice of condensed metallic hydrogen possessed tremendous similarity with graphite [19]. Could the layered form of metallic hydrogen finally replace Stoney's solid carbon on the Sun [2, 7]? Was this the strange material sought by Hastings for generating the solar spectrum [2, 8]?

2 Metallic hydrogen

Eugene Wigner (1963 Nobel Prize in Physics [21]) and Hillard B. Huntington [22] were the first to advance the existence of metallic hydrogen in 1935 [19]. They opened their classic paper by stating that "*Any lattice in which the hydrogen atoms would be translationally identical (Bravais lattice) would have metallic properties*" [19]. Their work focused on the body centered lattice. Recognizing the difficulties in obtaining the pressures required to form this lattice, they proposed that the layered form of metallic hydrogen would be more accessible. According to Wigner and Huntington "*it was J. D. Bernal who first put forward the view that all substances go over under very high pressure into metallic or valence lattices*" [19]. For the body centered cubic form of metallic hydrogen, they predicted a density of 0.8 g/cm^3 versus 0.087 g/cm^3 for molecular hydrogen in solid form [19]. This was nearly a tenfold increase in density. Wigner and Huntington concluded their paper as follows: "*The objection comes up naturally that we have calculated the energy of a body-centered metallic lattice only, and that another metallic lattice may be much more stable. We feel that the objection is justified. Of course it is not to be expected that another simple lattice, like the face-centered one, have a much lower energy, — the energy differences between forms are always very small. It is possible, however, that a layer-like lattice has a much greater heat of formation, and is obtainable under high pressure. This is suggested by the fact that in most cases of Table I of allotropic modifications, one of the lattices is layer-like*[19]*...*" [19]. The footnote in the text began: "*Diamond is a valence lattice, but graphite is a layer lattice...*" [19]. Thus, in the first paper on metallic hydrogen, the layered structure of graphite (see Figure 1), so critical to producing the blackbody spectrum on Earth, was promoted. A solar spectrum explained through dense hydrogen was certain to eventually rise to prominence.

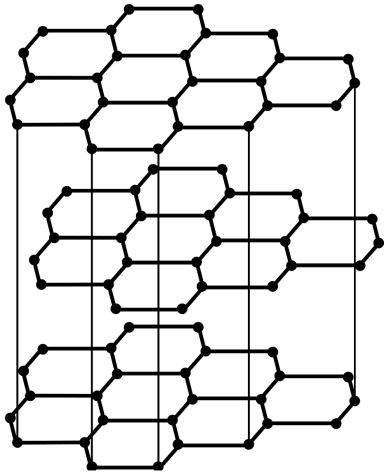


Fig. 1: Schematic representation of the layered lattice of graphite. Wigner and Huntington [19] would propose that most energetically favorable form of metallic hydrogen would assume this crystal structure.

2.1 Properties of metallic hydrogen

Initially, Wigner and Huntington estimated that the metallic state of hydrogen, in its most energetically accessible form (layered lattice), could be achieved at pressures in the 250,000 atm range (~25 GPa) [19]. This value was much too optimistic.

The most astounding property of metallic hydrogen would be its tremendous critical temperature. It was well in excess of anything Thomas Andrews and his contemporaries

could have imagined in 1869 [3,4]. While the complete phase diagram for hydrogen may never be fully known, several attempts have been made to outline its general characteristics, both in condensed matter physics [23–25] and as related to astrophysics [26–28]. Franck [29] listed many of the early contributions to the hydrogen phase diagram, including the work by Alexey A. Abrikosov [30]. Abrikosov eventually won the 2003 Nobel prize in physics while at Argonne National Laboratories.

The critical point of metallic hydrogen has been constantly revised towards ever higher values. Ebeling and Richert [23] provided an overview of these estimates through the 20th century. In 1980, Franck [29] arrived at a critical temperature for metallic hydrogen in the 6,000–9,000 K range. In 1983, Ronik and Kundt [26] gave a critical point at a unprecedented 19,100 K and 24 GPa. A slightly more conservative 16,500 K and 22.5 GPa was soon published [23]. Beyond critical temperatures, the transition pressures in moving from molecular to metallic hydrogen have constantly been revised upwards. At present, the values have moved to the 400–600 GPa range: "*Although quantum chemistry calculations have been developed to a high degree of sophistication, and in general, there is a close correlation between theory and experiment, this is not the case for hydrogen. Phase transition calculations that seek the structure with the lowest lattice energy have difficulty handling the zero-point energy contributions to the total energy and zero-point energy is very important for hydrogen. As a result, the predicted critical transition pressures have an enormous variation, from as low as 0.25 Mbar to over 20 Mbar, while recent predictions are in the 400 to 600 GPa range*" [25].

2.2 The theory of metallic hydrogen

Several authors have reviewed the metallic hydrogen literature [31, 32]. In a landmark 1968 publication, Neil Ashcroft hypothesized that metallic hydrogen might be a high temperature superconductor [33]. Ashcroft consequently became one of the most important theoretical physicists with respect to understanding dense hydrogen in its molecular and metallic forms [24, 33–50]. Ashcroft's prediction relative to high temperature superconductivity was rapidly echoed by Schneider and Stoll [51]. Depending on lattice configurations, they calculated that metallic hydrogen would become superconductive with operational temperatures ranging from 67 to 200 K [51]. Barbee et al. confirmed these calculations, obtaining a temperature of 230±85 K [52]. Metallic hydrogen had the potential to be the highest temperature superconductor known. The point was emphasized in 2001, when Maksimov and Savrasov used *ab initio* calculations to conclude that metallic hydrogen at high pressure might have a superconducting critical temperature of 600 K [53].

Ashcroft also examined the ground state of metallic hydrogen at zero temperature under conditions of changing spatial densities achieved by varying pressures from ∼1 Mbar to ∼75 Mbar [34, 35]. At the highest densities ($r_s = 0.8$, 1.2, 1.36, and 1.488), he discovered that crystalline phases were preferred [35]. However, at the lowest lattice density studied ($r_s = 1.64$), he found that metallic hydrogen was metastable between the solid and liquid forms [34, 35]. He postulated that the existence of a liquid ground state could not be excluded, but that it was not established [34]. Ashcroft continued this line of investigation in 1981 and 1982 [36, 37]. He gathered that liquid metallic hydrogen might become essentially devoid of structure and that the protons and electrons would simply act as interpenetrating fluids [36]. The Cornell scientist had theoretically constructed a two-component Fermi-liquid from protons and electrons [36].

Still, there was no direct evidence that metallic hydrogen at absolute zero would ever completely lose all structural integrity. As a theoretical physicist, Ashcroft could not really establish if metallic hydrogen at absolute zero 1) acted as a two component Fermi liquid, 2) behaved much like the unusual theoretical one-component plasma [54, 55], or 3) retained the essential characteristic of a Bravais lattice, an ordered proton field with fully degenerate electrons. Nonetheless, in his 1981 communication, Ashcroft was careful to mention that his conclusions were "*assuming that it [the hypothetical state of liquid metallic hydrogen] is normal*" [36]. He highlighted: "*that in assuming the existence of a liquid phase, the very interesting question still remains of whether or not it exhibits some form of magnetic, momental, or even spatial (e.g. liquid crystal) ordering... We do not attempt at this time to resolve the important questions of the existence or properties of possible "ordered" liquid metallic phases of hydrogen*" [36]. In the ninth footnote to his 1982 treatment, Ashcroft repeated the warning: "*The possibility that liquid metallic hydrogen exhibits some kind of momental (e.g. superconductive), magnetic, or even spatial (e.g. liquid crystal) ordering has not been ruled out*" [37]. Only experimental evidence could answer such questions, but none was available, as liquid metallic hydrogen remained an elusive material [25, 31, 32, 56].

Astrophysics was quick to infer that Ashcroft had chosen a path eventually leading to some form of degeneracy of matter [57]. In fact, careful reading of these articles suggested otherwise. Ashcroft's liquid was a reflection of what theoretical condensed matter physicists were able to calculate at the time. A liquid with spatial order, thoughtfully preserved in the text [36] and in the footnotes [37] of his papers, was well beyond the reach of computational approaches in the absence of laboratory guidance.

Soon after Ashcroft published his groundbreaking papers [34–37], MacDonald and Burgess also wondered about the absence of crystallization in metallic hydrogen [58]. They insisted that, since electronic screening was important in the solid state but negligible in the liquid state, metallic hydrogen would remain fluid at all pressures. Solid metallic hydrogen could not be stable at any pressure [58]. Ashcroft answered

that "*The prospect of a relatively low-density quantum melted phase of hydrogen, over a wide range of densities, is a fascinating one. However, we would like to bring up the following difficulties with concluding too definitely the existence of this phase for all densities*" [38]. Ashcroft then argued that such a state would exist only over limited densities whose range would be difficult to predict, as the solid and liquid phases are both close in energy and widely separated in configuration [38].

When Ashcroft returned to the ground state of metallic hydrogen in 1984, he assumed that the protons occupied the sites of a rigid Bravais lattice [39]. Using the Wigner-Seitz approximation which he regarded as physically appealing, Ashcroft calculated a lower bounds on the density of metallic hydrogen at its transition pressure. This density would be on the order 0.60 g/cm^3 corresponding to $r_s = 1.65$ [39]. Metallic hydrogen, if it was stable at all, would have to possess a greater density.

Given the nature of metallic hydrogen, both as a theoretical problem and as a prized material, significant Russian and Ukranian contributions were made in this area [32,53,59–65], beginning with Alexey Abrikosov [30]. In an important communication, Abrikosov was one of the first to examine the destruction of an atomic lattice under high compression [59]. He noted: "*that at sufficiently small volumes the positive zero-point oscillation energy exceeds the negative Coulomb energy, and this leads to a destruction of the crystal lattice*" [59]. Abrikosov remarked that "*the inter-atomic distances at the transition point are greater than the nuclear dimensions only for the lightest elements, hydrogen and helium. Thus, such a transition can take place only in these two elements*" [59]. It seemed as though elevated pressures might lead to the destruction of the crystal lattice, but Abrikosov never considered that fusion might act to relieve the stresses of compression. Beyond a certain point, perhaps crystals became incompressible. It was unclear if the small volumes required to give prominence to the zero-point oscillations in metallic hydrogen might ever be reached.

After Abrikosov's classic paper was released [59], Brovman et al. were the first to hypothesize that metallic hydrogen might be a metastable substance [61]. Kagan's group [32, 61] advanced that metallic hydrogen synthesized at elevated pressures might be completely stable even at zero pressure. This behavior would be much like diamond, the metastable form of carbon. Brovman et al. [61] calculated that the most stable lattice of metallic hydrogen would be hexagonal with a triangular string structure [60] . The conjecture would spawn the possibility of industrial and propellant roles for metallic hydrogen [25]. Many years later, Kaim et al. [64] would once again address the metastable nature of metallic hydrogen and essentially confirm Kagan's findings [61].

However, the most interesting facet of Kagan's work [61] was the observation that metallic hydrogen displayed liquid tendencies: "*there occurs in metallic hydrogen a unique tendency towards the formation of a family of structures with very close energies... In a certain sense the picture recalls the situation with graphite, but is apparently even more strongly pronounced... the formation of the planar family is evidence of the unique liquid-like tendencies that take place in metallic hydrogen under pressure*" [61]. They continued: "*As a result it is impossible to exclude beforehand, in principle, the possibility that the transition from the molecular phase to the metallic phase is a transition into the state of a liquid metal. (It may turn out that the situation will be different in hydrogen than in deuterium.) The phase diagram could have in this case a very special character. For example, with increasing pressure, the liquid phase could go over into the crystalline phase, but at extremely high densities a liquid would again be produced, but now as a result of the predominant role of the energy of the zero point oscillations (see the paper by Abrikosov[7]). The metastable state could remain crystalline in this case*" [61]. The footnote referred to the work just discussed above by Abrikosov [59]. Relative to the liquid metallic hydrogen model of the Sun, the work by Brovman et al. [61] would remain landmark.

Barbee et al. [52, 66] continued the quest to calculate the most stable structure for hydrogen in solid form. The work supported Wigner and Huntington's [19] contention that a layered Bravais lattice form of metallic hydrogen was the most stable in the 380±50 to 860±100 GPa range [66]. Above such values, the body centered cubic was preferred. Below 380±50 GPa the molecular non-metallic hexagonal-close-packed arrangement was most stable. The authors highlight some of the difficulties faced by theoretical condensed matter physics: "*A metal-insulator phase is expected near 200 GPa, in the m-hcp phase, but this transition pressure is harder to predict because of the shortcomings of local-density theory and the fact that structures with similar enthalpies (e.g. diamond and graphite) may have completely different band structures*" [66].

At about the same time, an interest developed in theoretical physics for examining the mono-, di-, and trilayered forms of atomic hydrogen [67–69]. While it could be argued that such structures were not physically realistic, their study generated additional insight into metallic hydrogen. Significantly, they demonstrated that very small changes in lattice parameters could alter the conductive behavior substantially, creating insulators from metals.

For his part, Neil Ashcroft maintained his interest in the structure of hydrogen. In 1993, he once again examined the metal-insulator transition in this element [40]. At this time, Ashcroft moved increasingly towards the idea that dense hydrogen might lack local structure at the lower densities. It seemed as if the stability of crystal forms was becoming questionable for him, even at the higher densities: "*At sufficiently high densities ($r_s \leqslant 1.5$), the predicted states of H (eq. 1) certainly include monatomic crystalline arrangements [6], at least where the dynamics of the protons can be ignored*" [40].

Though recognizing the presence of crystalline forms, he emphasized the dynamics of the protons. Observing that the proton pairing in molecular hydrogen was robust, Ashcroft eventually proposed that molecular metallic hydrogen might be energetically preferred [42]. This was a material very different than first proposed by Wigner and Huntington [19]. At low temperatures and at pressures less than 110 GPa, Ashcroft argued that molecular hydrogen existed as a rotational crystal [40,42]. At low densities ($1.5 < r_s < 2$), he envisioned that hydrogen might become a low temperature quantum fluid [45].

Ashcroft moved further towards the idea that, at the proper density, liquid hydrogen was a superfluid [47, 48]. In doing so, he revisited the ideas elucidated when first dealing with two component Fermi liquids [36, 37] and expanded on his work with Moulopoulos [41]. Ashcroft appropriately highlighted that experiments up to 300 GPa proved that molecular H-H stretching modes continued to exist at these high pressures [47]. He insisted that both proton-proton and electron-electron pairing could become the dominant interaction, given the proper conditions [41]. The concept that liquid metallic hydrogen was a two gap superconductor was also promoted by Babaev [70]. In such a superfluid, both protons and electrons could flow in the same direction, providing mass transfer without charge transfer. Alternatively, the system could result in a superconducting mode wherein proton and electrons flowed in opposite direction, resulting in the flow of both mass and charge [48]. Ashcroft then emphasized that "*the neutral superfluid mode does not couple to an external magnetic field, while the charged superconducting mode does*" [48]. The work did not address metallic hydrogen in its densest form. Ashcroft mentions that: "*Above any superconducting transition temperature (and above any Bose condensation temperature) liquid metallic hydrogen and deuterium should begin to adopt properties similar to those of conventional liquid metals, at least in the structural characteristics important to electron scattering*" [47].

Ashcroft's hypothesis that metallic hydrogen might exist as a quantum fluid immediately gained theoretical support [71]. Given increasing compression, Bonev et al. [71] calculated that solid molecular hydrogen [72] would be transformed into a quantum fluid state. Additional pressure would then lead to the monatomic crystal [19, 71]. With increasing pressure, it could be computed that hydrogen might undergo a transition from a liquid-molecular state into a non-molecular liquid [71]. This would become known as the liquid-liquid transition [71]. By extending the work of Brovman et al. [61], it was possible to visualize that hydrogen had a zero-temperature structured liquid ground state. With enough pressure, hydrogen could then move from the two component Fermi liquid [36, 37, 41, 46–48, 70], to the crystalline solid [19], and finally into a zero-temperature structured liquid state [61]. Alternatively, metallic hydrogen might move from a two component Fermi system directly either into a structured liquid metal [61] or into the solid classical form of metallic hydrogen [19]. A wide array of theoretical possibilities now existed for the state of hydrogen under dense conditions.

While the theory of liquid metals [73] has remained a fascinating branch of condensed matter physics, hydrogen liquid metals, though they appear simple on the surface, continued to offer unequalled challenges. With only sparse experimental data (see Section 2.4), theoretical condensed matter physics had little guidance from the laboratory. Even so, progress was being made, if only in the realization that metallic hydrogen was a material filled with mystery and promise. Modern condensed matter theory persisted in providing exciting results, often from the most prestigious groups [74–81].

Relative to solar physics, it was clear that the superfluid form of metallic hydrogen [36,37,41,46–48,70], devoid of all structure, could never be found on the surface of the Sun. The material required a very specific critical density along with low temperatures not found on the solar surface. Superfluid metallic hydrogen resembled nothing of the layered structure [19, 61] which mimicked graphite and was most likely to generate the solar spectrum. Superfluid metallic hydrogen [36, 37, 41, 46–48, 70] might never be found anywhere.

Fillinov et al. [74] studied dense hydrogen states at temperatures ranging from 10,000 to 100,000 K examining plasma phase transitions. Interestingly, at 10,000 K, they noticed droplet formation at certain densities (10^{23} cm^{-3}). But at the highest densities studied (10^{26} cm^{-3}), they observed an ordering of protons into a Wigner crystal. These were tremendous densities on the order of $\sim$150 g/cm^3. Militzer and Graham extended theoretical calculations to the petapascal range, a full eight orders of magnitude beyond the pressures of the molecular phase [76]. Such computations were appropriate only for the interior of astrophysical objects. Militzer and Graham [76] considered astounding hydrogen densities (2100 g/cm^3), but, in contrast to Abrikosov classic paper [59], the lattice was not destroyed and the calculations open serious questions as to the nature of the solid state.

Remaining in the realm of physically attainable pressures, Attaccalite and Sorella [77] demonstrated that the molecular liquid phase of hydrogen should be stable at pressures on the order of 300 GPa at $\sim$400 K. The melting curves for hydrogen and its phase boundaries have likewise been addressed [78, 79] revealing that theoretical approaches have remained difficult and open to new discoveries. Miguel Morales, while working with David Ceperley and Carlo Pierleoni [80], recently addressed the problem of metallic hydrogen by considering a range of temperatures and densities ($2,000 \leqslant T \leqslant 10,000$ K; $0.7 \leqslant \rho \leqslant 2.4$ g/cm^{-3}). Such conditions were appropriate for liquid metallic hydrogen devoid of structure, much like the one-component plasma [54, 55]. At elevated temperatures and densities, the system was observed to be a fully metallic liquid plasma [80]. However, a combination of lower densities and temperatures resulted in formation of an insulator [80]. Ceperley's group also con-

sidered electrical conductivity in high pressure liquid metallic hydrogen [81]. The work was noteworthy, as it tried to examine the liquid-semiconductor to liquid metal transition first reported experimentally by Weir et al. [82, 83]. Using either 32 or 54 atom cells, they calculated the transition density to be near $r_s \sim 1.65$, a value very close to the experimentally determined number ($r_s = 1.62$) [82, 83]. These calculations assumed that the liquid was devoid of any structure. In addition, David Ceperley examined hydrogen at ultra high pressures, $P \geqslant 20$ TPa [84], a value considerably lower than that of Militzer and Graham [76]. Furthermore, the Urbana-Champaign scientist studied the phase diagram for hydrogen in the ground state [85]. However, the theoretical procedure utilized was best suited to tritium and deuterium, as infinitely massive protons were hypothesized to be present. This work presented an excellent literature review and a remarkable array of potentially significant new structures for the ground state of hydrogen as a function of increasing pressure up to 5 TPa [85].

2.3 Metallic hydrogen in astrophysics

Soon after Wigner and Huntington [19] published their classic paper, liquid metallic hydrogen entered the realm of astrophysics. Its introduction as a constituent of the giant planets and the white dwarfs far preceded any experimental confirmation. Liquid metallic hydrogen would eventually occupy a peripheral position in astronomy, well removed from the Sun and most stars of the main sequence.

In 1946, Kronig et al. [86] proposed that metallic hydrogen existed at the center of the Earth. Their work was motivated by a recent report postulating that the Earth's center was composed of residual solar matter containing up to 30% hydrogen. Kronig et al. [86] calculated a density for metallic hydrogen of 0.8 g/cm^3. The result was apparently independent of Wigner and Huntington [19] as they seemed unaware of this previous communication. Then in 1950, W.H. Ramsey extended the study of metallic hydrogen to the planets and the white dwarfs [87]. According to Ramsey, at the International Astronomical Union meeting in Paris of 1935, H. N. Russell [88] had pointed out: "*that both the planets and the white dwarfs are cold in the sense that the density at any interval point is determined by the pressure at that point. In other words, the influence of temperature is so small that it can be neglected to a good approximation. Thus, in the accepted theory of the white dwarfs it is assumed that the electrons constitute a degenerate Fermi gas at absolute zero temperature*" [87]. The minutes of the meeting highlight how Russell believed that the maximum radius of a cold body was equal to one tenth of the solar radius, or about the diameter of Jupiter [88, p. 260]. It was a crucial statement which linked studies of the giant planets with those of the white dwarfs. At the pressures inside white dwarfs and giant planets, all solids were viewed as metallic [87]. Hydrogen was no exception. In the end, Ramsey deduced that metallic hydrogen could not be produced inside a small planet like the Earth [87]. Hence, it was primarily because of this work [87] that the quest for liquid metallic hydrogen would be extended simultaneously to the celestial objects with features of mass and density lying to either side of the Sun. In these objects, the study of liquid metallic hydrogen [26–28, 89, 90] progressed quickly to the fully degenerate liquid state (i.e. — states where both protons and electrons were unrestricted by lattice confinements).

Astrophysical bodies are not pure laboratory samples. They are an assembly of mixtures and alloys. As such, once scientists gained interest into the composition of the planets [91–95] and the white dwarfs (see [96] for a short review relative to ^{22}Ne), hydrogen/helium mixtures [97,98] and their alloys [49, 50, 99] were certain to attract attention. Along with Ashcroft, Eva Zurek and her coworkers [50] discovered that lithium had the capability of greatly stabilizing the metallization of hydrogen. Even the phase diagram for carbon under extreme conditions grew in importance, as potentially relevant to understanding Neptune, Uranus, and the white dwarf [100, 101]. A vast number of publications flourished, but they shared one common factor: the paucity of laboratory data. Nellis et al. extended results from the laboratory to interior of Jupiter [94, 95], well before his findings [82] were independently confirmed. Nellis' work on the production of liquid metallic hydrogen (see Section 2.4) at 140 GPa and 3,000 K was supported by conductivity measurements [82], although the merits of these measurements were to remain in doubt. In any case, astrophysics continued to insist that the large planets and white dwarfs were constituted of liquid metallic hydrogen devoid of structure and existing in fully degenerate states. At pressure of ~500 GPa (5 Mbar), William Nellis maintained that materials were either semiconductors or fully degenerate metals [102]. Experimental confirmation of a fully degenerate state for liquid metallic hydrogen at such pressures was unproven. In the laboratory, all forms of metallic hydrogen remained ethereal with theoretical predictions far surpassing experimental reality.

2.4 Laboratory quests for metallic hydrogen

Throughout the 20th century, the study of extraordinary states of matter has represented one of the most fascinating aspects of physics [102, 103]. The generation of extremes in temperatures, pressures, and densities has always involved complex and sophisticated experimental resources, often attainable only through national or multinational initiatives [103]. Nonetheless, with regards to metallic hydrogen [102], many efforts have been conducted in university level laboratories. Frederic Golden has provided an excellent review of the search for metallic hydrogen which Ho-Kwang Mao dubbed the "Holy Grail" of condensed matter physics [104]. Golden touches on the early Russian and American attempts to synthesize the material, along with a general description of methods [104]. Given the prize [56], experimental progress has been limited.

In June 1989, Ho-Kwang Mao and Russell Hemley, from the Geophysical Laboratory of the Carnegie Institution, reported evidence of metallization for hydrogen at 77 K and 250 GPa in the journal *Science* [105]. The key finding was the near opaqueness of the sample at the highest pressures. Isaac Silvera, working at the Lyman Laboratory of Physics at Harvard, was studying the metallization problem in parallel with Hemley [106–111]. He rapidly contested the validity of Hemley's claims and submitted a letter to Science [106] to which Mao and Hemley responded [107]. Silvera argued that visual darkening provided insufficient evidence for metallization and that further tests were needed [106]. Mao and Hemley defended their result, but in the end, conceded that "*The observations and spectroscopic measurements clearly indicate that significant changes in solid hydrogen occur with increasing pressure, but further work is needed to characterize in detail its optical, electrical, and structural properties under these conditions*" [107]. Silvera soon reported that there was no evidence of metallization up to 230 GPa from 77 to 295 K [110]. Metallic hydrogen had slipped away, but Ho-Kwang Mao, Russell Hemley, and Isaac Silvera would come to rank amongst the experimental leaders in the struggle to synthesize the material.

A few years later, Weir, Mitchell, and Nellis reported anew that metallic hydrogen had been produced [82]. Using shock compressed experiments [102, p. 1510–1514], the metallization of fluid molecular hydrogen was thought to have been achieved at 140 GPa and 3,000 K [82]. The communication was supported through conductivity measurements [82] a vital link in establishing metallization. The results were once again contested [112], though Nellis and Weir maintained their position [113]. In arguing against metallization, Besson brought in data with deuterium suggesting that its samples might represent highly degenerate material, something very different from molecular metallization in hydrogen [112]. Beyond this, Besson was concerned that the Al_2O_3 windows had affected the experiment [112]. Nellis and Weir countered that "*Our experiment and analysis yield the simple picture of a dense metallic fluid comprised primarily of molecular H_2 dimers and a relatively low dissociation fraction of ∼5% of H monomers*" [115]. The entire sequence of observation was on the order of just a few hundred nanoseconds [102, p. 1512], hardly time to conduct detailed structural analysis, while introducing tremendous difficulties in properly measuring both pressures and conductivities. William Nellis once again addressed his metallization experiments, but this time with Neil Ashcroft as a coauthor [114]. During the discussion which followed the paper, Nellis admitted that "*the exact nature of this unusual fluid needs to be determined*" [114, p. 135]. Though Nellis eventually claimed that "*Metallic fluid H is readily produced by dynamic high pressures*" [102, p. 1564], only questionable evidence existed for this state [82]. The shock experiments of metallic hydrogen from this group produced no additional results and other groups never confirmed the findings. The lack of lattice structure was debatable and mankind was no closer to metallic hydrogen. For his part, William Nellis moved to arguments of degeneracy, without solid experimental grounds [102].

In 1996, a collaboration between the University of Paris and the Geophysical Laboratory at the Carnegie Institution would make the next vital step forward [115]. Loubeyre et al. [115] examined both solid hydrogen and deuterium with X-ray diffraction at pressures just exceeding 100 GPa at 300 K. They discovered that solid hydrogen "*becomes increasingly anisotropic with pressures*" [115]. In like manner, the layered structure of graphite was considered anisotropic. Loubeyre et al. [115] tried to generate the equation of state for hydrogen as a function of temperature and pressure. They concluded that their results differed substantially from ab initio calculations "*indicating that theoretical understanding of the behavior of dense hydrogen remains incomplete*" [115]. Narayana et al. then studied solid hydrogen up to 342 GPa at 300 K [116]. These were pressures similar to those at the center of the Earth [117], but no evidence of metallization was found. The findings confirmed Ramsey's conclusion that the interior of the Earth could not support the metallic state of hydrogen [87]. In 2002, Loubeyre et al. again presented evidence that solid hydrogen became black, this time at 320 GPa and 100 K [118]. These values were not far removed from the 250 GPa used by Mao and Hemley in 1989 [105]. By observing the vibron mode, they maintain that molecular hydrogen in the solid form existed at least until 316 GPa, but Narayana had just reported that solid hydrogen remained transparent up to 342 GPa at 300 K [116]. Two of the world's major groups were again at odds with one another. Perhaps the discrepancies could be explained by difficulties in recording proper pressures at such values [102, p. 1514–1533]. After all, these studies were far from trivial in nature. Loubeyre et al. [118] refrained from stating that metallization had been achieved. Rather, they predicted that the process should occur near 450 GPa [118].

Mankind has remained unable to synthesize metallic hydrogen in the laboratory. However, as pressures rose and experimental settings improved, the characteristics of dense hydrogen did become increasingly established [119–125]. Great attention was placed on constructing phase diagrams for hydrogen (see [119] for a review). Determination of the peak in the melt line of this element has consequently been the subject of intense study (e.g. [121–124]). By this time, the broken symmetry and hydrogen-A phase for dense hydrogen were reasonably established, but neither form was metallic (see [122] for a brief review). Blackbody radiation finally entered such studies, with the goal to properly establish temperatures [122]. Along these lines, statements such as: "*we have shown that the emissivity of platinum is essentially independent of temperature in the temperature region of our study*" [122] would only serve as a reminder that not all was correct

with our understanding of blackbody radiation [13, 14]. For its part, metallic hydrogen continued to be ephemeral.

2.5 Commentary on liquid metallic hydrogen

As was seen in Section 2.3, within astrophysics, liquid metallic hydrogen is believed to exist as fully degenerate matter within the interior of white dwarfs and giant planets such as Jupiter or Saturn. Some have suggested that these planets also possessed liquid metallic helium, or a liquid metallic alloy of hydrogen and helium. Solid metallic hydrogen would have no role in astrophysics [27], as every hypothesis was either a molecular or a fully degenerate liquid. The conjecture that condensed matter could become degenerate in the large planets was far from what Chandrasekhar had envisioned when he first promoted degeneracy [57]. As a fully degenerate material, liquid metallic hydrogen could not sustain any useful current or magnetic field. Positive charges in liberal motion along with negative charges do not seem very amicable, either to potential generation or net current flow. At the same time, current flow with mass transfer seemed unreasonable in astrophysical objects. Direct laboratory observations remained much too elusive to reach any confirmation of these theoretical ideas. Some element of structure might always exist in metallic hydrogen independent of temperature. The superfluid form could remain ever theoretical, as Ashcroft had first carefully cautioned in the work with Oliva [36, 37].

The application of fully degenerate matter to the large planets and the white dwarfs was an unusual concept in light of a fully gaseous Sun. If Jupiter contained metallic hydrogen as degenerate matter and the same was true for the white dwarf, then it would not be unreasonable to place at least some condensed hydrogen on the Sun. Solar temperatures would prevent degenerate states and thus layered liquid metallic hydrogen represented a remarkable constitutive element.

When it was first conceived, the most energetically accessible form of metallic hydrogen was the layered lattice arrangement similar to that of graphite. Solid metallic hydrogen was viewed almost as a one component plasma [54, 55], wherein all electrons were degenerate and distributed over a hexagonal Bravais lattice formed from ordered protons [19]. In this sense, solid metallic hydrogen was considered as degenerate only relative to the flow of its electrons. Today, theoretical astrophysics has abandoned early thoughts of solid or liquid metallic hydrogen possessing a Bravais lattice [19], opting instead for fully degenerate materials where both protons and electrons flow freely. Conversely, experimentalists hope to harness metallic hydrogen for processes as varied as earthly fusion and rocket propulsion [25]. Such processes would not be easily approachable with a fully degenerate material. Hence, many experimental physicists are likely to be skeptical of a fully degenerate state for metallic hydrogen.

The progress towards dense hydrogen states has been an intriguing aspect of condensed matter physics. Ashcroft's two component Fermi liquid has remained a fascinating substance. However, given the combination of low temperatures, exact densities, and atypical conductive properties, it could have little practical role in human advancement. Current flow involving mass displacement was a concept which seemed to oppose structural stability, even though it could sustain magnetic fields. Conversely, when proton and electron displacement occurred in the same direction, there could be no current or the generation of magnetically interesting properties.

Theoretical condensed matter physics promoted hydrogen at extreme densities [76, 84], but hydrogen might not be compressible to such levels. In permitting essentially infinite compression of the lattice, it was debatable whether or not condensed matter physics had adopted a behavior similar to the ideal gas. Moreover, if compression was great enough, the solid might resist further attempts at reducing lattice dimensions. Fusion might relieve the stresses associated with compression.

3 Lessons from the Sun

Though the Sun would always remain devoid of the great advantage of our earthly laboratories, it has historically provided us with an amazing insight into nature. When Sir Joseph Lockyer and Pierre Jules César Janssen independently observed the lines of helium within solar spectra acquired in 1868 [126–130], they must have wondered if this unknown element would ever be discovered. Lockyer named this element Hēlios, the Greek name for the Sun god and the Sun [126]. Eventually, William Ramsay would isolate helium from cleveite [131–133], and the Sun would be credited for providing the first indication that helium existed. The identification of Coronium would follow a parallel story [134–136]. It took nearly three quarters of a century for Bengt Edlén and Walter Grotrian to finally identify Coronium from transition lines produced by highly oxidized iron, like Fe^{+13} and Fe^{+14} [136, p. 170]. Hence, a combination of earthly science and celestial observations became critical to the development of astronomy. This spirit of discovery has taught astronomers how to tackle even the most perplexing problems. The understanding of the solar spectrum should not be an exception.

3.1 Graphite, metallic hydrogen, and the solar spectrum

If graphite played a critical role, both in the construction of blackbodies [14], and historically in the structure of the Sun itself [2], it was because science has always recognized that graphite possessed a unique ability towards the production of Planck's spectrum [6, 13, 14]. Hastings was searching for a material which would possess many of the properties of graphite [8]. Graphite, the layered form of carbon, differed significantly in optical properties from its cubic counterpart, diamond. Structure was vital to the production of spectra. That materials were condensed was not sufficient, but a distinct lattice arrangement seemed central [9]. As a consequence, it would be expected that the layered form of metallic

hydrogen would resemble graphite itself in its optical properties. In contrast, fully degenerate forms of hydrogen [36, 37, 41, 46–48, 70–72] could never approach such optical behavior. Devoid of a true lattice, such a substance, if it truly existed anywhere, would be completely unable to generate a blackbody spectrum [6]. These are the lessons from our earthly laboratories, after examining thousands of materials over extreme ranges of frequencies and temperatures [13, 14, 137]. The structural lattice of graphite and soot was to remain unique in its thermal properties [13, 14]. It should serve as a guide for the nature of any condensed material placed either on the photosphere or within sunspots. The generation of a thermal spectrum with a blackbody lineshape has been solely a quality of condensed matter, not of gases, degenerate matter, or any other state which physicists might create.

Unlike the giant planets, the Sun possessed a unique feature: the ability to generate tremendous internal pressures and temperatures. Based on the solar spectrum [138–140] and other physical evidence [141], it was therefore reasonable to postulate that liquid metallic hydrogen must constitute the bulk of the solar mass and specifically the photospheric material [20, 142–149]. In considering a solar building block, thermal emission required a distinct lattice [150], as the absence of such structure would lead to the stellar opacity problem [9]. The author has previously made the point: "*As a result, the photosphere must be treated as condensed matter. Unfortunately, it is counterintuitive than an object at extreme temperatures can possess lattice structure. Nonetheless, given the evidence for condensed matter*[4]*, the solar constitutive element (primarily H) must form a lattice. The presence of powerful solar magnetic fields and gravitational forces make liquid metallic hydrogen a distinct possibility for the condensed state of the photosphere. In this case, the hydrogen nuclei can be viewed as arranged in an array forming an essentially incompressible solar lattice. The hydrogen electrons are contained within the metallic conduction bands. The inter-nuclear distance is being maintained by the need to keep the quantum conditions such that metallic conduction bands can be produced. Hydrogen contains no inner shell electrons. All the electrons are completely delocalized within the metallic conduction bands. As such, hydrogen in this state is not only a liquid metal (reminiscent of liquid sodium) but can also be viewed as a liquid metallic plasma*" [149]. The footnote referred to reference [141] in this work.

In the solar framework, the electrons would translate freely within the confines of conduction bands formed by the Bravais lattice of the protons. Though not a one-component plasma in a theoretical sense [54, 55], liquid metallic hydrogen could be considered as a one-component plasma in the physical sense since the electrons were delocalized. But liquid metallic hydrogen would possess a true Bravais lattice and, perhaps, even liquid crystal behavior [151–153]. In this regard, Ashcroft had left open the possibility that liquid metallic hydrogen was a liquid crystal in 1981 and 1982 [36, 37]. Ashcroft had been unable to exclude the possibility when he advanced the two-component Fermi liquid [36, 37]. Liquid metallic hydrogen could well have an ordered lattice which oscillates between structural forms. The finding by Brovman et al. [61] that metallic hydrogen, much like graphite, could adopt a family of structures with nearly the same energy should be considered in this regard.

In any event, it would be difficult to conceive that conduction bands could truly exist without a lattice and the importance of the Bravais lattice in the formation of metals should not be dismissed. To a large extent, liquid metallic hydrogen should preserve the layered structure of solid metallic hydrogen as anticipated by Wigner and Huntington [19]. But the metallic character might be somewhat reduced in the low pressures of the photosphere. In fact, this could be advantageous for emission, better resembling graphite. Indeed, if the graphitic spectrum was to be produced, the structure and conductive properties of liquid metallic hydrogen should resemble graphite as much as possible. This is because graphite represents the premier laboratory model.

3.2 Metallic hydrogen and solar structure

Metallic hydrogen, with its critical temperatures in the thousands of degrees Kelvin [23–26], overcomes all concerns raised regarding a liquid Sun based on Andrews [20] and his findings in ordinary gases [3, 4]. A liquid Sun composed of metallic hydrogen benefits from elevated critical temperatures for liquefaction, permitting hydrogen to adopt a condensed state even within an object like the Sun. Along these lines, it is doubtful that metallic hydrogen could really become infinitely compressed. Such a scenario appears unlikely, as the presence of conduction bands involves quantum restrictions on the lattice. If the internuclear distances are not ideal, quantum mechanical conditions should fail to support conduction. Two boundary conditions should exist. If the interatomic distance becomes too large, the substance should become an insulator. Similarly, if the interatomic distance becomes too small, the crystal should collapse [59] and conduction cease. In this respect, it would be important to note that the Sun has dynamo action and maintains large magnetic fields. Both of these phenomena make destruction of the conducting lattice unlikely [141].

It remains unclear why condensed structures resist compression, but invoking fusion as a means of releasing the strain of compressions should be a viable solution. This is especially the case if compared to the destruction of the crystal [59] and the creation of fully degenerate matter [36, 37, 41, 46–48, 70]. Degeneracy removes all of the forces which lead to fusion. As such, it should be more reasonable to maintain the relative incompressibility of condensed matter. The Sun, after all, has a very ordinary density of 1.4 g/cm^3 [141] and the same is true for the giant planets. Thus, Jeans' idea that the Sun represents a rotating liquid mass of reasonably constant density should not be dismissed [2]. Condensed mat-

ter and metallic hydrogen provide a framework for ordinary densities, even in light of enormous pressures. The reward of such an approach is threefold leading to: 1) a reasonable framework to generate the solar spectrum, 2) a decent ability to impart structure, and 3) a practical path towards fusion.

A Sun composed of metallic hydrogen provides an interesting model to explain sunspots and other structural elements. The photospheric material in this case might be considered as liquid metallic hydrogen where the lattice dimensions are relaxed at lowered pressures. Perhaps, the material exists much like graphite at the limits of conductive behavior. Conversely, within sunspots, pressures would be more elevated, and liquid metallic hydrogen might assume a more compact lattice, with increased metallic behavior. This would help account for the stronger magnetic fields observed within sunspots. As a result, scientists could be considering the conversion from a Type I lattice in the photosphere to a Type II lattice in the sunspots [141]. Such a scenario has great advantages in terms of simplicity.

Gases have always been an unsustainable building material for an object like the Sun. Gases know no surface and cannot, even momentarily, impart structure. Hence, one cannot be surprised to find that there is no physical evidence which supports a gaseous Sun, while ample evidence [141] has been revealed for its condensed state [20, 142–149]. In order to bring structure to the gas, astrophysics must depend on the action of magnetic fields. However, strong magnetic fields themselves are a property of condensed matter, not gases [141]. In order to maintain a gaseous Sun and impart it with structures, astrophysics must therefore have recourse to phenomena best produced by condensed matter.

A simple illustration of these issues can be focused on the understanding of solar prominences. Such objects appear as sheet-like structures in images captured by NASA's SOHO satellite (see Figure 2). In a Sun built from layered metallic hydrogen, it can be envisioned that a layer of material simply peeled away from the surface to form a prominence. In contrast, within a gaseous body, the creation of such overwhelming structures would remain difficult to explain, even with magnetic fields forming and maintaining these entities. Perhaps it would be more logical to presume that magnetic fields were simply associated with the presence of metallic hydrogen, whether on the surface of the Sun itself or within the prominences.

Moreover, the active photosphere and chromosphere supports structural features [154]. Prominences contain fine structure [155, 156], which would be easier to explain if a condensed solar model was adopted. For more than one century [157, p. 104], prominences have been known to emit continuous spectra in addition to the line spectra which characterize the quiescent state [158–161]. Eilnar Tandberg-Hanssen has long studied prominences and has provided an excellent review of the subject matter [160]. Like other solar physicists, because the Sun was considered as a gas, he viewed promi-

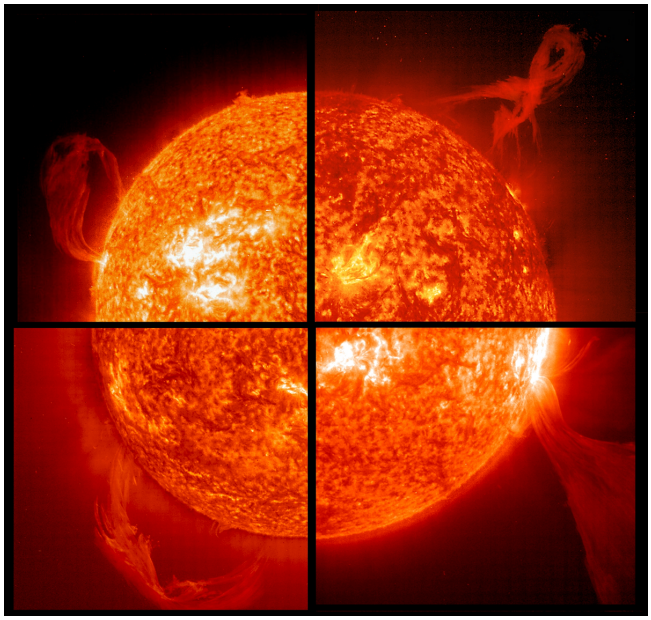


Fig. 2: Sheet like appearance of solar prominences. NASA describes the image as follows: "*A collage of prominences, which are huge clouds of relatively cool dense plasma suspended in the Sun's hot, thin corona. At times, they can erupt, escaping the Sun's atmosphere. For all four images, emission in this spectral line of EIT 304Å shows the upper chromosphere at a temperature of about 60,000 degrees K. The hottest areas appear almost white, while the darker red areas indicate cooler temperatures. Going clockwise from the upper left, the images are from: 15 May 2001; 28 March 2000; 18 January 2000, and 2 February 2001*". Courtesy of SOHO/[Extreme ultraviolet Imaging Telescope (EIT)] consortium. SOHO is a project of international cooperation between ESA and NASA. http://sohowww.nascom.nasa.gov/gallery/images/promquad.html (accessed May 31, 2011).

nences as gaseous in nature [160]. Tandberg-Hanssen maintained that the continuous spectrum associated with some quiescent prominences was being generated by the scattering of light emitted from the photosphere [161]. This was because gaseous prominences could have no means of generating continuous spectra by themselves. They should have produced only line spectra. Conversely, if the Sun was made from condensed metallic hydrogen, the prominences could directly produce the continuous spectrum. No scattering would need to be invoked. If the density of the prominence material in some cases could not sustain a continuous spectrum, then only line spectra would be generated. Thus, as the prominence dissipated with time, it would be expected that the continuous spectrum might weaken or become absent. It is possible to consider that prominences are formed by layered metallic hydrogen separating from the inferior levels of the photosphere. A slight change in density could account for such actions reflecting an abrupt transformation from a more compact lattice to a less dense form. This hypothesis might explain why entire sheets of material appear to be ejected, some-

thing which would be difficult to understand otherwise.

It is possible, one further observation worth pondering involves a figure presented by Fortov in his new text [103]. The figure in question (Figure 7.7 in [103]) consists of a plot of the log of object diameter versus the log of mass. On such a plot, a straight line passes through all astrophysical objects within our solar system, from the smallest comic dust, to the meteorites, to the comets, to the asteroids, to the satellites of planets, to the planets, and finally to the Sun [103, p. 192]. This plot provides another line of evidence that the Sun should be viewed as condensed matter. Every object on the graph can be considered as condensed. Uranus and Neptune are currently viewed as having metallosilicate cores and mantles of ices [103, p. 193]. Jupiter and Saturn are largely liquid metallic hydrogen or helium in either molecular or atomic form [103, p. 193]. As the only remaining fully gaseous object in the solar system, it may be reasonable to suggest that the Sun should not stand alone on such a graph.

4 Conclusion

Relative to the Sun, a condensed approach brings interesting contrasts and dilemmas versus the gaseous models. The latter are endowed with tremendous mathematical flexibility [1, 2], but their physical relevance appears limited. Gases cannot by themselves impart structure and the solar spectrum is not easily explained in a gaseous framework [9]. The gaseous stars suffer from the stellar opacity problem [9]. Conversely, a liquid metallic hydrogen model imparts a wonderful ability to explain the origin of the solar spectrum relying on the layered structure held in common with graphite [141–149]. Metallic hydrogen possesses a very high critical temperature and can exist as condensed matter even on the solar surface accounting for many features of the Sun best characterized by material endowed with a lattice [141]. Most of the physical attributes of the Sun are more simply explained within the framework of a liquid model [141]. However, a condensed Sun is not as open to theoretical formulations. The advantages of a liquid Sun are now so numerous [20, 141–149] that it is difficult to conceive why the model was not proposed long ago. This speaks to the allure of the gaseous Sun and the mathematical beauty of the associated equations of state.

In closing, it should be highlighted that there is currently an effort to describe the Sun as "liquid-like" (e.g. [162]). In the end, the author believes that such terminology should be avoided. If the Sun is condensed, it should be viewed as liquid, not "liquid-like". Even gases could be "liquid-like". Such terms cannot be sufficient, since a real lattice is required for production of the solar thermal spectrum. No compromise can be made on this point for those who have studied thermal emission in real materials. "Liquid-like" might refer to anything from a gas, to a plasma, to fully degenerate matter, to supercritical fluid and none are necessarily endowed with a lattice. The contention of this work remains that the photosphere of the Sun is liquid, with *true lattice structure and ordered interatomic distances*. The adoption of liquid metallic hydrogen as a solar constituent brings with it a wealth of possibilities in describing solar structures and understanding the solar spectrum. Central to this advancement, the lattice must remain the foremost element in all of condensed matter, whether here on Earth, within the Sun, and even, in the firmament of the stars.

Acknowledgement

Luc Robitaille is acknowledged for producing a rendition of graphite's layered lattice.

Dedication

This work is dedicated to my son, Christophe, and his wife, Lindsey.

Submitted on May 31, 2011 / Accepted on June 07, 2011
First published online on June 9, 2011

References

1. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v.3, 3–25.
2. Robitaille P.M. A thermodynamic history of the solar constitution — II: The theory of a gaseous sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v.3, 41–59.
3. Andrews T. The Bakerian lecture: On the continuity of the gaseous and liquid states of matter. *Phil. Trans. Roy. Soc.*, 1869, v.159, 575–590.
4. Rowlinson J.S. Thomas Andrews and the critical point. *Nature*, 1969, v.224, 541–543.
5. Fison A.H. Recent Advances in Astronomy. Blackie and Son, London, 1898.
6. Planck M. Ueber das Gesetz der Energieverteilung in Normalspectrum. *Annalen der Physik*, 1901, v.4, 553–563.
7. Stoney G.J. On the physical constitution of the sun and stars. *Proc. Roy. Soc. London*, 1867, v.16, 25–34.
8. Hastings C.S. A theory of the constitution of the Sun, founded upon spectroscopic observations, original and other. *Am. J. Science*, 1881, v.21(121), 33–44.
9. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v.3 — a paper published in this Special Issue.
10. Schuster A. Radiation through a foggy atmosphere. *Astrophys. J.*, 1905, v.21, 1–22.
11. Watts W.M., Huggins W. An Introduction to the Study of Spectrum Analysis. Longmans, Green, and Co., London, 1904.
12. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsverögen und dem Absorptionsvermögen. der Körper für Wärme und Licht. *Poggendorff's 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.*, ser. 4, 1860, v.20, 1–21).
13. Robitaille P.M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v.4, 3–13.
14. Robitaille P.M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v.3, 36–55.
15. Jeans J.H. Problems of Cosmogony and Stellar Dynamics — Being an essay to which the Adams Prize of the University of Cambridge for the year 1917 was Adjudged. Cambridge University Press, 1919.

16. Payne C.H. The relative abundances of the elements, in: Stellar Atmospheres, Harvard University Monograph no. 1, Harvard University Press, 1925, chapter 13 (Also found in Lang K.R. and Gingerich O.: A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p. 244–248).
17. Unsöld A. Über die Struktur der Fraunhofersehen Linien und die quantitative Spektralanalyse der Sonnenatmosphäre. *Zeitschrift für Physik*, 1928, v.46, 765–781.
18. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v.70, 11–82.
19. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v.3, 764–770.
20. Robitaille P.M. Liquid metallic hydrogen: Building block of a liquid Sun. *APS Ohio Spring Meeting*, 2011, D4.00005.
21. Telegdi V. Obituary: Professor Eugene Wigner. *The Independent*, Friday January 13, 1995.
22. Giaever I. and Grone A.R. Hillard B. Huntington. *Physics Today*, 1993, v.46(8), 70.
23. Ebeling W. and Richert W. Thermodynamic properties of liquid hydrogen metal. *Phys. Stat. Sol. B*, 1985, v.128, 467–474.
24. Edwards B. and Ashcroft N.W. Order in dense hydrogen at low temperatures. *Proc. Nat. Acad. Science USA*, 2004, v.101(12), 4013–4018.
25. Cole J.W. and Silvera I.F. Metallic hydrogen: The most powerful rocket fuel yet to exist. *J. Phys: Conference Series*, 2010, v.215, 012194(1–9).
26. Robnik M. and Kundt W. Hydrogen at high pressures and temperatures. *Astron. Astrophys.*, 1983, v.120, 227–233.
27. Hubbart W.B., Guillot T., Lunine J.I., Burrows A., Saumon D., Marley M.S. and Freedman R.S. Liquid metallic hydrogen and the structure of brown dwarfs and giant planets. *Phys. Plasmas*, 1997, v.4(5), 2011–2015.
28. Hubbart W.B., Burrows A. and Lunine J.I. Theory of giant planets. *Ann. Rev. Astron. Astrophys.*, 2002, v.40, 103–136.
29. Franck S. On the dielectric-metal transition in hydrogen. *Ann. Physik (Leipzig)*, 1980, v.492(5), 349–356.
30. Abrikosov A.A. Uravnenie sostoyaniya vodoroda pri vysokikh davleniyakh. *Astronomicheskii Zhurnal*, 1954, v.31, 112.
31. Mao H.K. and Hemley R.J. Ultrahigh-pressure transitions in solid hydrogen. *Rev. Mod. Physics*, 1994, v.66(2), 671–692.
32. Maksimov E.G. and Shilov Yu.I. Hydrogen at high pressure. *Physics Uspekhi*, 1999, v.42(11), 1121–1138.
33. Ashcroft N.W. Metallic hydrogen: A high-temperature superconductor? *Phys. Rev. Letters*, 1968, v.21(26), 1748–1749.
34. Chakravarty S. and Ashcroft N.W. Ground state of metallic hydrogen. *Phys. Rev. B*, 1978, v.18, 4588–4597.
35. Mon K.K., Chester G.V. and Ashcroft N.W. Simulation studies of a model of high-density metallic hydrogen. *Phys. Rev. B*, 1980, v.21, 2641–2646.
36. Oliva J. and Ashcroft N.W. Two-component Fermi-liquid theory: Equilibrium properties of liquid metallic hydrogen. *Phys. Rev. B*, 1981, v.23, 6399–6407.
37. Oliva J. and Ashcroft N.W. Two-component Fermi-liquid theory: Transport properties of liquid metallic hydrogen. *Phys. Rev. B*, 1982, v.25, 223–236.
38. Carlsson A.E. and Ashcroft N.W. Liquidlike state of dense hydrogen. *Phys. Rev. B*, 1984, v.29(1), 479–480.
39. Styer D.F. and Ashcroft N.W. Ground-state energy of metallic hydrogen in the Wigner-Seitz approximation. *Phys. Rev. B*, 1984, v.29(10), 5562–5569.
40. Ashcroft N.W. The metal-insulator transition: complexity in a simple system. *J. Non-Crystalline Solids*, 1993, v.156–158, 621–630.
41. Moulopoulos K. and Ashcroft N.W. Coulomb interactions and generalized pairing in condensed matter. *Phys. Rev. B*, 1999, v.59(19), 12309–12325.
42. Johnson K.A. and Ashcroft N.W. Structure and bandgap closure in dense hydrogen. *Nature*, 2000, v.403, 632–635.
43. Ashcroft N.W. The hydrogen liquids. *J. Phys: Condensed Matter*, 2000, v.12, A129–A137.
44. Nagao K., Bonev S.A. and Ashcroft N.W. Cusp-condition constraints and the thermodynamic properties of dense hot hydrogen. *Phys. Rev. B*, 2001, v.64, 224111(1–12).
45. Bonev S.A. and Ashcroft N.W. Hydrogen in jellium: First principles of pair interactions. *Phys. Rev. B*, 2001, v.64, 224112(1–9).
46. Ashcroft N.W. Hydrogen at high density. *J. Phys. A: Math. Gen.*, 2007, v.36, 6137–6147.
47. Nagao K., Bonev S.A., Bergars A., and Ashcroft N.W. Enhanced Friedel structure and proton pairing in dense solid hydrogen. *Phys. Rev. Letters*, 2003, v.90(3), 035501(1–4).
48. Ashcroft N.W. Observability of a projected new state of matter: a metallic superfluid. *Phys. Rev. Letters*, 2005, v.95(10), 105301(1–4).
49. Ashcroft N.W. Hydrogen dominant metallic alloys: High temperature superconductors? *Phys. Rev. Letters*, 2004, v.92(18), 187002(1–4).
50. Zurek E., Hoffmann R., Ashcroft N.W., Oganov A.R., Lyakhov A.O. A little bit of lithium goes a lot for hydrogen. *Proc. Nat. Acad. Sci. USA*, 2009, v.106(42), 17640–17643.
51. Schneider T. and Stoll E. Metallic hydrogen II high-temperature superconductivity. *Physica*, 1971, v.55, 702–710.
52. Barbee T.W., Garcia A. and Cohen M.L. First-principles prediction of high-temperature superconductivity in metallic hydrogen. *Nature*, 1989, v.340, 369–371.
53. Maksimov E.G. and Savrasov D.Yu. Lattice stability and superconductivity of metallic hydrogen at high pressure. *Solid State Communications*, 2001, v.119, 569–572.
54. Jones M.D. and Ceperly D.M. Crystallization of the one-component plasma at finite temperature. *Phys. Rev. Letters*, 1996, v.76(24), 4572–4575.
55. Daligault J. Liquid-state properties of a one-component plasma. *Phys. Rev. Letters*, 2006, v.96, 065003(1–4); and erratum *Phys. Rev. Letters*, 2009, v.103, 029901.
56. Pool R. The chase continues for metallic hydrogen. *Science*, 1990, v.247(4950), 1545–1546.
57. Chandresekhar S. The highly collapsed configurations of a stellar mass — Second paper. *Mon. Not. Roy. Astron. Soc.*, 1935, v.95, 207–225.
58. McDonald A.H. and Burgess C.P. Absence of crystallization in metallic hydrogen. *Phys. Rev. B*, v.26(6), 2849–2855.
59. Abrikosov A.A. Some properties of strongly compressed matter — I. *Sov. Phys. JETP*, 1961, v.12(6), 1254–1259.
60. Brovman E.G., Kagan Yu. and Kholas A. Structure of metallic hydrogen at zero pressure. *Sov. Phys. JETP*, 1972, v.34(6), 1300–1315.
61. Brovman E.G., Kagan Yu.M. and Kholas A. Properties of metallic hydrogen under pressure. *Sov. Phys. JETP*, 1972, v.35(4), 783–787.
62. Kagan Yu.M., Pushkarev V.V., Kholas A. Equation of state of the metallic phase of hydrogen. *Sov. Phys. JETP*, 1977, v.46(3), 511–522.
63. Petrov Yu.V. Molecular hydrogen crystal in the multiparticle approximation. *Sov. Phys. JETP*, 1983, v.57(2), 449–452.
64. Kaim S.D., Kovalenko N.P., Vasiliu E.V. Many particle interactions and local structure of the metallic hydrogen at zero pressure. *J. Phys. Studies*, 1997, v.1(4), 589–595.

65. Belashchenko D.K. The simulation of metallic hydrogen-helium solutions under the conditions of internal Jupiter regions. *Russian J. Phys. Chem. A*, *Focus on Chemistry*, 2006, v.80, S31–S39.
66. Barbee T.W., Garcia A., Cohen M.L. and Martins J.L. Theory of high-pressure phases of hydrogen. *Phys. Rev. Letters*, 1989, v.62, 1150–1153.
67. Wu J.Z., Sabin J.R., Trickey S.B. and Boettger J.C. Mono- and di-layer analogues of crystalline atomic hydrogen. *Int. J. Quant. Chem.: Quantum Chemistry Symposium*, 1990, v.24, 873–879.
68. Nobel J.A., Wilson G.A. and Trickey S.B. Near-equilibrium ordering of the crystalline phase of metallic hydrogen. *Int. J. Quant. Chem.*, 1992, v.42, 1037–1045.
69. Wu J.Z., Trickey S.B., Sabin J.R. and Boettger J.C. Structure, energetics, and molecular- to atomic-ordering transitions in hydrogen thin films. *Phys. Rev. B*, 1992, v.45, 8610–8622.
70. Babaev E. Vortices with fractional flux in two-gap superconductors and in extended Faddeev model. *Phys. Rev. Letters*, 2002, v.89(6), 067001(1–4).
71. Bonev S.A., Schwegler E., Ogitsu T. and Galli G. A quantum fluid of metallic hydrogen suggested by first-principles calculations. *Nature*, 2004, v.431, 669–672.
72. Kranendonk J.V. Solid Hydrogen: Theory of the Properties of Solid H_2, HD, and D_2. Plenum Press, New York, 1983.
73. March N.H. Liquid Metals: Concepts and Theory. Cambridge University Press, Cambridge, 1990.
74. Filinov V.S., Bonitz M., Fortov V.E., Ebeling W., Lavshov P. and Schlanges M. Thermodynamic properties and plasma phase transition in dense hydrogen. *Contrib. Plasma Phys.*, 2004, v.44(5–6), 388–394.
75. Delaney K.T., Pierleoni C. and Ceperley D.M. Quantum Monte Carlo simulation of the high-pressure molecular-atomic crossover in fluid hydrogen. *Phys. Rev. Letters*, 2006, v.97, 235702(1–4).
76. Militzer B. and Graham R.L. Simulations of dense atomic hydrogen in the Wigner crystal phase. *J. Phys. Chem. Solids*, 2006, v.67, 2136–2143.
77. Attaccalite C. and Sorella S. Stable liquid hydrogen at high pressure by a novel ab initio molecular-dynamics calculation. *Phys. Rev. Letters*, 2008, v.100, 114501(1–4).
78. Davis S.M., Belonoshko A.B., Johansson B., Skorodumova N.V., and van Duin A.C.T. High pressure melting curve of hydrogen. *J. Chem. Phys.*, 2008, v.129, 194508(1–5).
79. Tamblyn I. and Bonev S.A. Structure and phase boundaries of compressed liquid hydrogen. *Phys. Rev. Letters*, 2010, v.104, 065702(1–4).
80. Morales M.A., Pierleoni C. and Ceperley D.M. Equation of state of metallic hydrogen from coupled electron-ion Monte Carlo simulations. *Phys. Rev. E*, 2010, v.81, 021202(1–9).
81. Lin F., Morales M.A., Delaney K.T., Pierleoni C., Martin R.M., and Ceperley D.M. Electrical conductivity of high-pressure liquid hydrogen by quantum Monte Carlo methods. *Phys. Rev. Letters*, 2009, v.103, 256401(1–4).
82. Weir S.T., Mitchell A.C. and Nellis W.J. Metallization of fluid molecular hydrogen at 140 GPa (1.4 Mbar). *Phys. Rev. Letters*, 1996, v.76(11), 1860–1863.
83. Nellis W.J., Weir S.T. and Mitchell A.C. Minimum metallic conductivity of fluid hydrogen at 140 GPa (1.4 Mbar). *Phys. Rev. B*, 1999, v.59, 3434–3449.
84. Liberatore E., Pierleoni C. and Ceperley D.M. Liquid-solid transition in fully ionized hydrogen at ultra high-pressures. *J. Chem. Phys.*, 2011, v.134, 184505(1–11).
85. McMahon J.M. and Ceperley D.M. Ground-state structures of atomic metallic hydrogen. *Phys. Rev. Letters*, 2011, v.106, 165302(1–4).
86. Kronig R., de Boer J. and Korringa J. On the internal constitution of the Earth. *Physica*, 1946, v.12(5), 245–256.
87. Ramsey W.H. The planets and the white dwarfs. *Mon. Not. Roy. Astron. Soc.*, 1950, v.110, 444–454.
88. Unsigned. The International Astronomical Union meeting in Paris 1935. *The Observatory*, 1935, v.58, 257–265.
89. Salpeter E.E. Energy and pressure of a zero-temperature plasma. *Astrophys. J.*, 1961, v.134(3), 669–682.
90. Hubbard W.B. and Slattery W.L. Statistical mechanics of light elements at high pressure I. Theory and results for metallic hydrogen with simple screening. *Astrophys. J.*, 1971, v.168, 131–139.
91. Brown H. On the composition and structures of the planets. *Astrophys. J.*, 1950, v.111, 641–653.
92. Hubbard W.B. and Smoluchowski R. Structure of Jupiter and Saturn. *Space Science Rev.*, 1973, v.14(5), 599–662.
93. Gillan M.J., Alfe D., Brodholt J., Vocadlo L. and Price G.D. First principles modelling of Earth and planetary materials at high pressures and temperatures. *Rep. Prog. Phys.*, 2006, v.69, 2365–2441.
94. Nellis W.J., Ross M. and Holmes N.C. Temperature measurements of shock-compressed hydrogen: Implications for the interior of Jupiter. *Science*, 1995, 269(5228), 1249–1252.
95. Nellis W.J., Weir S.T. and Mitchell A.C. Metallization and electrical conductivity of hydrogen in Jupiter. *Science*, 1996, 273(5277), 936–938.
96. Deloye C. J. and Bildstein L. Gravitational settling of ^{22}Ne in liquid white dwarf interiors: cooling and seismological effects. *Astrophys. J.*, 2002, v.580, 1077–1090.
97. Stevenson D.J. Thermodynamics and phase separation of dense fully ionized hydrogen-helium mixtures. *Phys. Rev. B*, 1975, v.12(10), 3999–4007.
98. Vorberger J., Tamblyn I., Militzer B. and Bonev S.A. Hydrogen-helium mixtures in the interior of giant planets. *Phys. Rev. B*, 2007, v.75, 024206(1–11).
99. Hafner J. Structure and thermodynamics of liquid metals and alloys. *Phys. Rev. A*, 1977, v.16(1), 351–364.
100. Correa A.A., Bonev S.A. and Galli G. Carbon under extreme conditions: Phase boundaries and electronic properties from first-principles theory. *Proc. Nat. Acad. Sciences USA*, 2006, v.103(5), 1204–1208.
101. Silvera I. Diamond: Molten under pressure. *Nature Physics*, 2010, v.6, 9–10.
102. Nellis W.J. Dynamic compression of materials: metallization of fluid hydrogen at high pressures. *Rep. Prog. Phys.*, 2006, v.69, 1479–1580.
103. Fortov V.E. Extreme States of Matter on Earth and in the Cosmos. Springer-Verlag, Berlin, 2011.
104. Golden F. Metallic hydrogen: Another milestone in the solid state. *Mosaic*, 1991, v.22(2), 22–31.
105. Mao H.K. and Hemley R.J. Optical studies of hydrogen above 200 Gigapascals: Evidence for metallization by band overlap. *Science*, 1989, v.244, 1462–1464.
106. Silvera I.F. Evidence for band overlap metallization of hydrogen. *Science*, 1990, v.247(4944), 863.
107. Mao H.K. and Hemley R.J. Evidence for band overlap metallization of hydrogen — Response. *Science*, 1990, v.247(4944), 863–864.
108. Lorenzana H.E., Silvera I.F. and Goettel K.A. Evidence for a structural phase transition in solid hydrogen at megabar pressures. *Phys. Rev. Letters*, 1989, v.63, 2080–2083.
109. Lorenzana H.E., Silvera I.F. and Goettel K.A. Order parameter and a critical point on the megabar-pressure hydrogen-A phase line. *Phys. Rev. Letters*, 1990, v.65, 1901–1904.

110. Eggert J.H., Moshary F., Evans W.J., Lorenzana H.E., Goettel K.A. and Silvera I.F. Absorption and reflectance in hydrogen up to 230 GPa: Implications for metallization. *Phys. Rev. Letters*, 1991, v.66, 193–196.

111. Kaxiras E., Broughton J. and Hemley R.J. Onset of metallization and related transitions in solid hydrogen. *Phys. Rev. Letters*, 1991, v.67, 1138–1141.

112. Besson J.M. Comment on "Metallization of fluid molecular hydrogen at 140 GPa (1.4 Mbar)". *Phys. Rev. Letters*, 1997, v.78(26), 5026.

113. Nellis W.J. and Weir S.T. Nellis and Weir reply. *Phys. Rev. Letters*, 1997, v.78(26), 5027.

114. Nellis W.J., Louis A.A., and Ashcroft N.W. Metallization of fluid hydrogen. *Phil. Trans. Roy. Soc. London*, 1998, v.356, 119–138.

115. Loubeyre P., LeToullec R., Hausermann D., Hanfland M., Hemley R.J., Mao H.K. and Finger L.W. X-ray diffraction and equation of state of hydrogen at megabar pressures. *Nature*, 1996, v.383, 702–704.

116. Narayana C., Luo H., Orloff J. and Ruoff A.L. Solid hydrogen at 342 GPa: no evidence for an alkali metal. *Nature*, 1998, v.393, 46–49.

117. Unsigned. High-pressure scientists "journey" to the center of the Earth, but can't find elusive metallic hydrogen. *Cornell News*, May 6, 1998.

118. Loubeyre P., Occelli F. and LeToullec R. Optical studies of solid hydrogen to 320 GPa and evidence for black hydrogen. *Nature*, 2002, v.416, 613–617.

119. Kohanoff J. The status of low-temperature phase diagram of hydrogen at the turn of the century. *J. Low Temp. Phys.*, 2001, 122(3/4), 297–311.

120. Gregoryanz E., Goncharov A.F., Matsuishi K., Mao H.K. and Hemley R.J. Raman spectroscopy of hot dense hydrogen. *Phys. Rev. Letters*, 2003, v.90(17), 175701(1–4).

121. Matsuishi K., Gregoryanz E., Mao H.K. and Hemley R.J. Equation of state and intermolecular interactions in fluid hydrogen from Brillouin scattering at high pressures and temperatures. *J. Chem. Phys.*, 2003, 118(23), 10683–10695.

122. Deemyad S. and Silvera F. Melting line of hydrogen at high pressures. *Phys. Rev. Letters*, 2008, v.100, 155701(1–4).

123. Grinenko A., Gericke D.O., Glenzer S.H. and Vorberger J. Probing the hydrogen melt line at high pressures by dynamic compression. *Phys. Rev. Letters*, 2008, v.101, 194801(1–4).

124. Eremets M.I. and Trojan I.A. Evidence of maximum in the melting curve of hydrogen at megabar pressures. *JETP Letters*, 2009, v.89(4), 174–179.

125. Tolédano P., Katzke H., Goncharov A.F. and Hemley R.J. Symmetry breaking in dense solid hydrogen: Mechanism for the transitions to phase II and phase III. *Phys. Rev. Letters*, 2009, v.103, 105301(1–4).

126. Jensen W.B. Why Helium Ends in "-ium". *J. Chem. Educ.*, 2004, v.81, 944.

127. Lockyer J.N. Notice of an Observation of the Spectrum of a solar prominence (letter of October 20, 1868). *Phil. Mag.*, 1869, v.37, 143 (also found in Meadows A.J. Early Solar Physics. Pergamon Press, Oxford, 1970, p. 119).

128. Janssen P.J.C. The Total Solar Eclipse of August 1868, Part I. *Astron. Reg.*, 1869, v.7, 107–110.

129. Janssen P.J.C. Summary of some of the results obtained — The total solar eclipse of August 1868 — Part II. *Astron. Reg.*, 1869, v.7, 131–134.

130. Janssen P.J.C. Summary of some of the results obtained at Cocanada, during the eclipse last August, and afterwards. A letter from P.J.C. Janssen to the permanent secretary (found in Meadows A.J. Early Solar Physics. Pergamon Press, Oxford, 1970, p. 117–118).

131. Ramsay W. On a gas showing the spectrum of helium, the reputed cause of D3, one of the lines in the coronal spectrum. Preliminary note. *Proc. Roy. Soc. London*, 1895, v.58, 65–67.

132. Ramsay W. Helium, a gaseous constituent of certain minerals. Part I. *Proc. Roy. Soc. London*, 1895, v.58, 80–89.

133. Ramsay W. Helium, a gaseous constituent of certain minerals. Part II. *Proc. Roy. Soc. London*, 1895, v.59, 325–330.

134. Claridge G.C. Coronium. *J. Roy. Astron. Soc. Canada*, 1937, v.31, 337–346.

135. Unsigned. Origin of the coronium lines. *Nature*, 1942, v.150(3817), 756–759.

136. Curtis L.J. Atomic Structure and Lifetimes: a Conceptual Approach. Atomic Structure and Lifetimes: A Conceptual Approach. Cambridge University Press, Cambridge, 2003.

137. Touloukian Y.S. and Ho C.Y. Thermophysical Properties of Matter (vols. 1–8). Plenum, New York, 1970.

138. Langley S.P. Experimental determination of wave-lengths in the invisible spectrum. *Mem. Natl. Acad. Sci.*, 1883, v.2, 147–162.

139. Langley S.P. On hitherto unrecognized wave-lengths. *Phil. Mag.*, 1886, v.22, 149–173.

140. Langley S.P. The invisible solar and lunar spectrum. *Am. J. Science*, 1888, v.36(216), 397–410.

141. Robitaille P.M. The solar photosphere: evidence for condensed matter. *Progr. Phys.*, 2006, v.2, 17–21 (also found in slightly modified form within *Research Disclosure*, 2006, v.501, 31–34; title #501019).

142. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v.1, 70–81 (also in arXiv: astro-ph/0410075).

143. Robitaille P.M. The collapse of the Big Bang and the gaseous Sun. *New York Times*, March 17, 2002, page A10 (available online: http://thermalphysics.org/pdf/times.pdf).

144. Robitaille P.M. Evidence for a liquid plasma model of the Sun. *Am. Phys. Soc. Meeting*, *April 2004*, S280002.

145. Robitaille P.M. The Sun as a hot liquid plasma: additional evidence. *Am. Phys. Soc. Meeting*, *Ohio Spring 2004*, S50002.

146. Robitaille P.M. The photosphere as condensed matter. *Am. Phys. Soc. Meeting*, *Ohio Fall 2004*, S60005.

147. Robitaille P.M. The Sun as a hot liquid plasma: more evidence. *Am. Phys. Soc. Meeting*, *NE Fall 2004*, S10004.

148. Robitaille P.-M.L. A radically different point of view on the CMB. In: *Questions of Modern Cosmology — Galileo's Legacy*, ed. by M. D'Onofrio and C. Burigana, Springer, New York, 2009.

149. Robitaille P.M. The Sun as a high energy/high density liquid metallic hydrogen plasma. *The 33rd IEEE International Conference on Plasma Science*, June 4–8, 2006, Traverse City, Michigan, p. 461, DOI:10.1109/PLASMA.2006.1707334.

150. Robitaille P.M.L. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v.31(6), 1263–1267.

151. de Gennes P.G. and Prost J. The Physics of Liquid Crystals. 2nd Edition, Oxford Science Publications, Oxford, 1993.

152. Chandrasekhar S. Liquid Crystals. Cambridge University Press, Cambridge,1992.

153. Mingos D.M.P. Liquid Crystals. Springer, Berlin, 1999.

154. Tsiropoula G. Structures and flows in the solar active photosphere and chromosphere. *2nd Advances in Solar Physics Euroconference*, *Three-Dimensional Structure of Solar Active Regions*, *ASP Conference Series*, 1998, v.155, 24–43.

155. Heinzel P. The fine structure of solar prominences. *The Physics of Chromospheric Plasmas*, *ASP Conference Series*, 2007, v.368, 271–291.

156. Gunár S., Heinzel P., Schmieder B., Schwartz P. and Anzer U. Properties of prominence fine-structure threads derived from SOHO/SUMER hydrogen Lyman lines. *Astron. Astrophys.*, 2007, v.472, 929–936.

157. Campbell W.W. The total solar eclipse of august 30, 1905. *The Popular Science Monthly*, 1904, v.65, 97–108.

158. Jefferies J.T. and Orrall F.Q. On the interpretation of prominence spectra. III. The line and continuous spectrum of a loop prominence and limb flare. *Astrophys. J.*, 1961, v.133, 963–969.

159. Yakovkin N.A. and Zeldina M.Yu. The prominence radiation theory. *Solar Physics*, 1975, v.45, 319–338.

160. Tandberg-Hanssen E. Solar prominences — An intriguing phenomenon. *Solar Physics*, 2011, v.269, 237–251.

161. Tandberg-Hanssen E. A spectroscopic study of quiescent prominences. *Astrophysica Norvegica*, 1964, v.9(3), 13–32.

162. Brown D.S. Our explosive sun. *Physics Education*, 2009, v.44 (1), 20–26.

On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye

Pierre-Marie Robitaille
Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA
E-mail: robitaille.1@osu.edu

In this exposition, the existence of the solar surface will be briefly explored. Within the context of modern solar theory, the Sun cannot have a distinct surface. Gases are incapable of supporting such structures. The loss of a defined solar surface occurred in 1865 and can be directly attributed to Hervé Faye (Faye H. Sur la constitution physique du soleil. *Les Mondes*, 1865, v.7, 293–306). Modern theory has echoed Faye affirming the absence of this vital structural element. Conversely, experimental evidence firmly supports that the Sun does indeed possess a surface. For nearly 150 years, astronomy has chosen to disregard direct observational evidence in favor of theoretical models.

Herbert Spencer was the first to advance that the body of the Sun was gaseous [1], but he believed, much like Gustav Kirchhoff [2], that the photosphere was liquid [3, 4]. For his part, Father Angelo Secchi [5, 6] promoted the idea that the Sun was a gaseous body with solid or liquid particulate matter floating within its photosphere. Soon after Father Secchi's second Italian paper [6] was translated into French by l'Abbé François Moigno, Harvé Faye made claims of independent and simultaneous discovery [3, 7, 8].

Harvé Faye almost immediately published his own work in *Les Mondes* [9]. In this communication, he deprived the Sun of its distinct surface. He based the loss of a solar surface on the gaseous nature of the interior and the associated convection currents. The salient sections of Faye's classic 1865 article stated: "*So then the exterior surface of the Sun, which from far appears so perfectly spherical, is no longer a layered surface in the mathematical sense of the word. The surfaces, rigorously made up of layers, correspond to a state of equilibrium that does not exist in the Sun, since the ascending and descending currents reign there perpetually from the interior to the superficial area; but since these currents only act in the vertical direction, the equilibrium is also not troubled in that sense, that is to say, perpendicularly to the leveled layers that would form if the currents came to cease. If, therefore, the mass was not animated by a movement of rotation, (for now we will make of it an abstraction), there would not be at its heart any lateral movement, no transfer of matter in the perpendicular direction of the rays. The exterior surface of the photosphere being the limit that will attain the ascending currents which carry the phenomenon of incandescence in the superior layers, a very-admissible symmetry suffices in a globe where the most complete homogeneity must have freely established itself, to give to this limit surface the shape of a sphere, but a sphere that is incredibly uneven*" [9].

In the same article, Hervé Faye emphasized that the photospheric surface was illusionary: "*This limit is in any case only apparent: the general milieu where the photosphere is incessantly forming surpasses without doubt, more or less, the highest crests or summits of the incandescent clouds, but we do not know the effective limit; the only thing that one is permitted to affirm, is that these invisible layers, to which the name atmosphere does not seem to me applicable, would not be able to attain a height of 3′, the excess of the perihelion distance of the great comet of 1843 on the radius of the photosphere*" [9]. Though astronomy has denied the existence of a distinct solar surface as a question of utmost complexity involving opacity arguments [10], the conjecture was actually proposed by Faye in 1865 within a framework of questionable value [9]. Hervé Faye's contributions to solar theory have been extensively addressed [3] and many, like his famous *Les Mondes* communication [9], were not supported by mathematics. Early solar theory rested on vague hypotheses.

It was only much later that Faye's ideas would gain the support of mathematical formulation. In 1891, August Schmidt of Stuttgart wrote a small pamphlet which solidified Faye's conjectures [11]. Within two years, Schmidt received the support of Knoft and, in 1895, Wilczynski published a detailed summary of their ideas in the *Astrophysical Journal* [11]. The illusionary nature of the solar surface was finally supported by mathematics. James Keeler was the first to voice an objection to Schmidt's theory, responding immediately to Wilczynski's article [12]: "*But however difficult it may be for present theories to account for the tenuity of the solar atmosphere immediately above the photosphere, and however readily the same fact may be accounted for by the theory of Schmidt, it is certain that the observer who has studied the structure of the Sun's surface, and particularly the aspect of the spots and other markings as they approach the limb, must feel convinced that these forms actually occur at practically the same level, that is, that the photosphere is an actual and not an optical surface. Hence it is, no doubt, that the theory is apt to be more favorably regarded by mathematicians than by observers*" [12]. Twenty years after Schmidt proposed his ideas, they had still not gained the support of observational astronomers such as Charles Abbot, the director of the Smithsonian Observatory: "*Schmidt's views have obtained considerable acceptance, but not from observers of solar phenomena*" [13, p. 232].

In 1896, Edwin B. Frost [14] discussed Wilson's theory [15] in which sunspots represented depressions on the solar photosphere [3]. He maintained that the theory was not yet well established and required further investigation. Nonetheless, the highlight of his paper would be a comment relative to the existence of a true solar surface. Frost's work [14] formed an appropriate reminder that the presence of the solar surface had been long denied by those who, by advocating gaseous solar models, must reject solar structure as mere illusion: "*In speaking of levels we must proceed from some accepted plane of reference; and the most natural plane, or surface of reference, would be the solar photosphere. Here we are abruptly confronted by the theory of Schmidt, elaborated in a convenient form by Knoft, according to which the photosphere is merely an optical illusion, produced by circular refraction in the Sun itself, supposed to be a globe of glowing gas without a condensed stratum. Prominences, faculae, spots, and granulation are explained as effects of anomalous refractions due to local changes of density somewhere in the gas ball. This theory, worked out as it is by careful mathematical reasoning, deserves and has received respectful consideration. Nevertheless, in view of the physical improbability of Schmidt's primary assumption that in its outer portions the gaseous mass maintains its state without condensation, the physicist will feel obliged to reject the theory, which also suffers from the fundamental defect of failing to account for the solar spectrum on the accepted principles of physics. Moreover, any one who has with some continuousness studied the phenomena of the solar surface must affirm that he has observed realities, not illusions. The perspective effects on prominences as they pass around the limb, the motion and permanence of the spots, the displacements of the spectral lines on the approaching and receding limbs, and in fact all the phenomena concerned with solar rotation, are distinctly contradictory to Schmidt's theory. In dismissing it from further consideration, however, we shall take with us the important inference that refraction within and on the Sun itself may modify in some considerable degree the phenomena we observed*" [14].

Though Faye and Schmidt denied the presence of a distinct surface on the Sun, it was clear that observational astronomers were not all in agreement. The point was also made in 1913 by Edward Walter Maunder, the great solar physicist: "*But under ordinary conditions, we do not see the chromosphere itself, but look down through it on the photosphere, or general radiating surface. This, to the eye, certainly looks like a definite shell, but some theorists have been so impressed with the difficulty of conceiving that a gaseous body like the Sun could, under the conditions of such stupendous temperatures as there exist, have any defined limit at all, that they deny that what we see on the Sun is a real boundary, and argue that it only appears so to us through the effects of the anomalous refraction or dispersion of light. Such theories introduce difficulties greater and more numerous than those that they clear away, and they are not generally accepted by the practical observers of the Sun. They seem incompatible with the apparent structure of the photosphere, which is everywhere made up of a complicated mottling: minute grains somewhat resembling those of rice in shape, of intense brightness, and irregularly scattered. This mottling is sometimes coarsely, sometimes finely textured; in some regions it is sharp and well defined, in others misty or blurred, and in both cases they are often arranged in large elaborate patterns, the figures of the pattern sometimes extending for a hundred thousand miles or more in any direction. The rice like grains or granules of which these figures are built up, and the darker pores between them, are, on the other hand, comparatively small, and do not, on the average, exceed two to four hundred miles in diameter*" [16, p. 28].

That same year, Alfred Fowler [17] the British spectroscopist who trained as Lockyer's assistant, commented on problems in astronomy [18]. Fowler served as the first secretary of the International Astronomical Union [17]. Fowler's writings reflected that the ideas of Hervé Faye [9] and August Schmidt [11] continued to impact astronomy beyond 1913 [3, 4], even though observational astronomers were not convinced: "*The apparently definite bounding-surface of the Sun which is ordinarily revealed to the naked eye, or seen in the telescope, has such an appearance of reality that its existence has been taken for granted in most of the attempts which have been made to interpret solar phenomena... Thus the photosphere is usually regarded as a stratum of cirrus or cumulus clouds, consisting of small solid or liquid particles, radiating light and heat in virtue of their state of incandescence... An effort to escape from this difficulty was made in the view suggested by Johnstone Stoney, and vigorously advocated by Sir Robert Ball, that the photospheric particles consist of highly refractory substances carbon and silicon (with a preference for carbon), both of which are known to exist on the Sun...The photosphere is thus regarded as an optical illusion, and remarkable consequences in relation to spots and other phenomena are involved. The hypothesis appears to take no account of absorption, and, while of a certain mathematical interest, it seems to have but little application to the actual Sun*" [18]. It was well known that Johnstone Stoney [19] advocated that the solar photosphere contained carbon particles [4].

Even in the 21st century, astronomy has maintained that the Sun's surface is an illusion. For instance, in 2003, the National Solar Observatory claimed that "*The density decreases with distance from the surface until light at last can travel freely and thus gives the illusion of a 'visible surface'*" [20].

Nonetheless, spectacular images of the solar surface have been acquired in recent years, all of which manifest phenomenal structural elements on or near the solar surface. High resolution images acquired by the Swedish Solar Telescope [20–23] reveal a solar surface in three dimensions filled with structural elements. Figure 1 displays an image which is publicly available for reproduction obtained by the Swedish So-

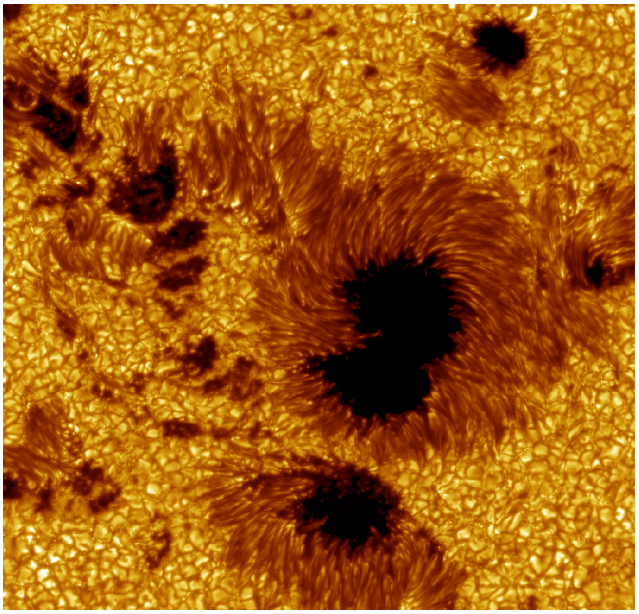


Fig. 1: Part of a sunspot group near the disk center acquired with the Swedish 1-m Solar Telescope by Göran B. Scharmer, Boris V. Gudiksen, Dan Kiselman, Mats G. Löfdahl, and Luc H. M. Rouppe van der Voort [21]. The image has been described as follows by the Institute for Solar Research of the Royal Swedish Academy of Sciences: "*Large field-of-view image of sunspots in Active Region 10030 observed on 15 July 2002. The image has been colored yellow for aesthetic reasons*" http://www.solarphysics.kva.se

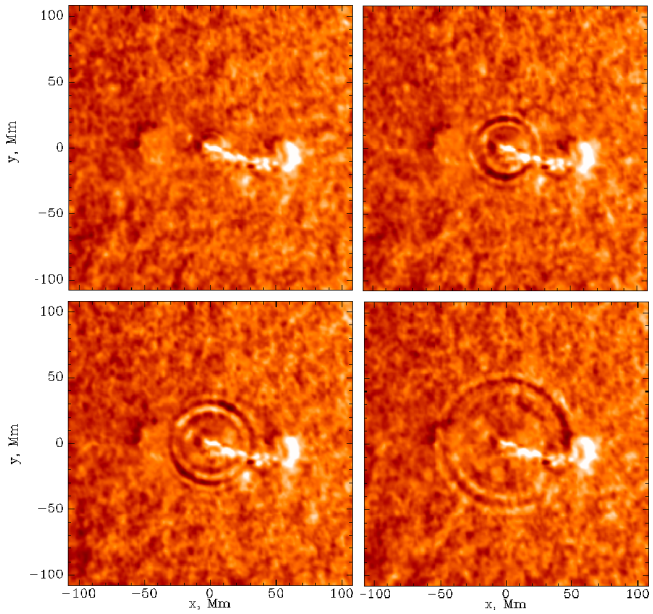


Fig. 2: Doppler image of a solar flare and the associated disturbance on the solar surface acquired by the NASA/ESA SOHO satellite. Such data was described as "*resembling ripples from a pebble, thrown into a pond*" [25]. Courtesy of SOHO/[Michelson Doppler Imager] consortium. SOHO is a project of international cooperation between ESA and NASA.

lar Telescope of the Royal Swedish Academy of Science. The author has previously commented on these results: "*The solar surface has recently been imaged in high resolution using the Swedish Solar Telescope [24, 25]. These images reveal a clear solar surface in 3D with valleys, canyons, and walls. Relative to these findings, the authors insist that a true surface is not being seen. Such statements are prompted by belief in the gaseous models of the Sun. The gaseous models cannot provide an adequate means for generating a real surface. Solar opacity arguments are advanced to caution the reader against interpretation that a real surface is being imaged. Nonetheless, a real surface is required by the liquid model. It appears that a real surface is being seen. Only our theoretical arguments seem to support our disbelief that a surface is present*" [24]. References [24] and [25] in the quotation referred to [21, 22] in the current work. A study of Lites et al. [23] illustrates how these authors hesitated to regard the solar surface as real, precisely because they considered that the Sun was gaseous in nature: "*However, since the angular resolution of the SST* [Swedish Solar Telescope] *is comparable to the optical scale of the photosphere (about one scale height), we may no longer regard the photospheric surface as a discontinuity; optical depth effects must be considered*" [23]. Though the authors reported three-dimensional structure, they added quotation marks around the word "surface" precisely because a gaseous Sun cannot support such a feature. They referred to the "optical depth unit surface", a concept inherently tied to gaseous models of the Sun. At the same time, the authors displayed a qualified desire for condensed matter: "*This gives the (perhaps false) visual impression of a solid surface of granules that protrude up a considerable distance from the surface, and that a raised structure is "illuminated" by a light source in the vicinity of the observer*" [23].

Beyond the evidence provided by the Swedish Solar Telescope and countless other observations, there was clear Doppler confirmation that the photosphere of the Sun was *behaving* as a distinct surface [25, 26]. In 1998, Kosovichev and Zharkova published their *Nature* paper *X-ray flare sparks quake inside the Sun* [25]. Doppler imaging revealed transverse waves on the surface of the Sun, as reproduced in Figure 2: "*We have also detected flare ripples, circular wave packets propagating from the flare and resembling ripples from a pebble, thrown into a pond*" [25]. In these images, the "optical illusion" *was now acting as a real surface*. The ripples were clearly transverse in nature, a phenomenon difficult to explain using a gaseous solar model. Ripples on a pond are characteristic of the liquid or solid state.

Hervé Faye's contention that the Sun was devoid of a real surface has never been supported by observational evidence; the solar surface has long ago been established. Though theory may hypothesize a gaseous Sun, it must nevertheless sup-

port observational findings. Perhaps, now that a reasonable alternative to a gaseous Sun has be formulated [27], astrophysics will discard the idea that the solar surface is an illusion, embrace the liquid nature of the Sun, and move to better comprehend this physical reality.

Acknowledgement

The Swedish 1-m Solar Telescope is operated on the island of La Palma by the Institute for Solar Physics of the Royal Swedish Academy of Sciences in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias.

Dedication

This work is dedicated to my eldest son, Jacob.

Submitted on June 9, 2011 / Accepted on June 13, 2011
First published online on June 24, 2011

References

1. Unsigned (Spencer H.) Recent astronomy, and the nebular hypothesis. *Westminster Review*, 1858, v.70, 185–225.
2. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H. E. Roscoe, Macmillan and Co., Cambridge, 1862, 23–31.
3. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v.3, 3–25 — a paper published in this Special Issue.
4. Robitaille P.M. A thermodynamic history of the solar constitution — II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v.3, 41–59 — a paper published in this Special Issue.
5. Secchi A. Sulla Teoria Delle Macchie Solari: Proposta dal sig. Kirchoff. *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 31 January 1864, v.3(4), 1–4 (translated into English by Eileen Reeves and Mary Posani: On the theory of sunspots proposed by Signor Kirchoff, *Progr. Phys.*, 2011, v.3, 26–29 — a paper published in this Special Issue).
6. Secchi A. Sulla Struttura della Fotosfera Solare. *Bullettino Meteorologico dell' Osservatorio del Collegio Romano*, 30 November 1864, v.3(11), 1–3. (translated into French by François Moigno: Sur la structure de la photosphère du soleil. *Les Mondes*, 22 December 1864, v.6, 703–707; translated into English by Mary Posani and Eileen Reeves: On the structure of the photosphere of the sun, *Progr. Phys.*, 2011, v.3, 30–32 — a paper published in this Special Issue).
7. Faye H.A.E.A. Sur la constitution physique du Soleil — première partie. *Comptes Rendus*, 1865, v.60, 89–96.
8. Faye H.A.E.A. Sur la constitution physique du Soleil – deuxième partie. *Comptes Rendus*, 1865, v.60, 138–150.
9. Faye H. Sur la constitution physique du Soleil. *Les Mondes*, 1865, v.7, 293–306 (translated into English by Patrice Robitaille: On the physical constitution of the Sun — Part I. *Progr. Phys.*, 2011, v.3, 35–40 — a paper published in this Special Issue).
10. Robitaille P.M. Solar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v.3 — a paper published in this Special Issue.
11. Wilczynski E.J. Schmidt's theory of the Sun. *Astrophys. J.*, 1895, v.1, 112–126.
12. Keeler J.E. Schmidt's theory of the Sun. *Astrophys. J.*, 1895, v.1, 178–179.
13. Abbot C.G. The Sun. D. Appleton and Company. New York, 1911.
14. Frost E.B. On the level of sun-spots. *Astrophys. J.*, 1896, v.4, 196–204.
15. Wilson A. Observations on the solar spots. *Phil. Trans. Roy. Soc.*, 1774, v.64, 1–30.
16. Maunder E.W. Are the planets inhabited? Harper & Brothers, London, 1913.
17. Unsigned. Obituary Notices: Fellows: Fowler, Alfred. *Mon. Not. Roy. Astron. Soc.*, v.101, 132-134.
18. Fowler A. Some problems in astronomy IV. Solar and stellar photospheres. *The Observatory*, 1913, v.36, 182–185.
19. Stoney G.J. On the physical constitution of the sun and stars. *Proc. Roy. Soc. London*, 1867, v.16, 25–34.
20. National Solar Observatory. Advanced Technology Solar Telescope — ATST. http://atst.nso.edu/files/press/ATST_book.pdf (accessed online on June 7, 2011).
21. Scharmer G. B., Gudiksen B.V., Kiselman D., Löfdahl M.G., Rouppe van der Voort L.H.M. Dark cores in sunspot penumbral filaments. *Nature*, 2002, v.420, 151–153.
22. MacRobert A. Solar faculae stand exposed. *Sky and Telescope*, 2003, v.106(4), 26.
23. Lites B.W., Scharmer G.B, Berger T.E., and Title A.M. Three-dimensional structure of the active region photosphere as revealed by high angular resolution. *Solar Physics*, 2004, v.221, 65–84.
24. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v.2, 17–21.
25. Kosovichev A. G., Zharkova V.V. X-ray flare sparks quake inside the Sun. *Nature*, 1998, v.393, 317–318.
26. Fleck B., Brekke P., Haugan S., Duarte L. S., Domingo V., Gurman J.B., Poland A.I. Four years of SOHO discoveries — some highlighs. *ESA Bulletin*, 2000, v.102, 68–86.
27. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v.3, 60–74 — a paper published in this Special Issue.

On Solar Granulations, Limb Darkening, and Sunspots: Brief Insights in Remembrance of Father Angelo Secchi

Pierre-Marie Robitaille

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

Father Angelo Secchi used the existence of solar granulation as a central line of reasoning when he advanced that the Sun was a gaseous body with a photosphere containing incandescent particulate matter (Secchi A. Sulla Struttura della Fotosfera Solare. *Bullettino Meteorologico dell'Osservatorio del Collegio Romano*, 30 November 1864, v.3(11), 1–3). Secchi saw the granules as condensed matter emitting the photospheric spectrum, while the darkened intergranular lanes conveyed the presence of a gaseous solar interior. Secchi also considered the nature of sunspots and limb darkening. In the context of modern solar models, opacity arguments currently account for the emissive properties of the photosphere. Optical depth is thought to explain limb darkening. Both temperature variations and magnetic fields are invoked to justify the weakened emissivities of sunspots, even though the presence of static magnetic fields in materials is not usually associated with modified emissivity. Conversely, within the context of a liquid metallic hydrogen solar model, the appearance of granules, limb darkening, and sunspots can be elegantly understood through the varying directional emissivity of condensed matter. A single explanation is applicable to all three phenomena. Granular contrast can be directly associated with the generation of limb darkening. Depending on size, granules can be analyzed by considering Kolmogoroff's formulations and Bénard convection, respectively, both of which were observed using incompressible liquids, not gases. Granules follow the 2-dimensional space filling laws of Aboav-Weiner and Lewis. Their adherence to these structural laws provides supportive evidence that the granular surface of the Sun represents elements which can only be constructed from condensed matter. A gaseous Sun cannot be confined to a 2-dimensional framework. Mesogranules, supergranules, and giant cells constitute additional entities which further support the idea of a condensed Sun. With respect to sunspots, the decrease in emissivity with increasing magnetic field strength lends powerful observational support to the idea that these structures are comprised of liquid metallic hydrogen. In this model, the inter-atomic lattice dimensions within sunspots are reduced. This increases the density and metallic character relative to photospheric material, while at the same time decreasing emissivity. Metals are well known to have lowered directional emissivities with respect to non-metals. Greater metallicity produces lower emissivity. The idea that density is increased within sunspots is supported by helioseismology. Thus, a liquid metallic hydrogen model brings with it many advantages in understanding both the emissivity of the solar surface and its vast array of structures. These realities reveal that Father Secchi, like Herbert Spencer and Gustav Kirchhoff, was correct in his insistence that condensed matter is present on the photosphere. Secchi and his contemporaries were well aware that gases are unable to impart the observed structure.

1 Introduction

The appearance of sunspots has fascinated mankind for centuries [1–8] and while limb darkening [9–11] has been documented from the days of Galileo [3, p.274], the phenomenon only became well-established in the 1800's [7, 12]. Solar granulations have also long captivated solar science [13, 14]. Although humanity has gazed at the Sun since time immemorial, our understanding of these phenomena remains limited. In a large measure, this reflects the unassailable nature of the Sun. At the same time, our lack of understanding mirrors the incapacity of the gaseous models to properly address questions related to solar structure. Gases will always remain devoid of structural attributes.

Strangely, if Father Angelo Secchi [2] first advanced that the Sun was constituted of a gaseous body surrounded by a photosphere containing particulate matter [16, 17], it was because he was searching to understand photospheric structure. The nature of solar granulations troubled Secchi [2, 17]. He solved the problem by endowing the body of the Sun with a gaseous nature while maintaining a partially condensed photosphere. Secchi's proposed photosphere could not adhere to the full properties of condensed matter. Sixty years later, theoretical physics advocated a completely gaseous solar model.

As a result, it has been nearly impossible to synthesize a realistic and cohesive portrayal of sunspots, granulation, and limb darkening, even though a cursory review of the question suggests otherwise.

2 Granulations and the gaseous models

2.1 Ideas of the 19th century

Secchi built his solar model on two driving forces: 1) Nasmyth's early description of solar granulation [18, 19] and 2) Magnus' demonstration that solid sodium hydroxide increased the luminosity of the gaseous flame [20]. Based on Magnus [20], Secchi advanced [17] that some condensed matter was present within the photosphere, as gases were devoid of the emissive power required to produce the solar spectrum [2]. Secchi considered that the darker appearance of intergranular lanes reflected the inferior radiative ability of the gaseous solar body. He believed that Nasmyth's discovery was noteworthy [18, 19], though remarking that granular features had previously been observed on the solar surface: "*First of all, are these new findings? We believe that, in the end, these are the same granulations that have long since been pointed out by observers, under the name of "lucules" and "pores" and that with the new method they can better be distinguished*" [17]. Secchi's description of granulation was important to the history of astronomy, as the Jesuit scientist was regarded as one of the leading solar observers of his time [2, 21]. His representations of granules depicted in his classic text [21, p.31–34] (reproduced in part within [14, p.4] and [1, p.143–145]) were nothing short of astounding. In 1870, Secchi presented drawings which remain respectable by today's standards and which far surpassed the illustrations which had made James Nasmyth famous only a few years before (see drawings reproduced in [13]).

In the mid-1860s, considerable controversy erupted between James Nasmyth [22] and the Reverend William Rutter Dawes [23] over the appearance of the solar granulation [13]. Nasmyth supported the notion that granules had a consistent structure and resembled regular overlapping "*willow leaves*". For his part, Dawes maintained that they had been discovered long before Nasmyth and that the term "*willow leaves*" was inappropriate as the features displayed an irregular form [13]. The discussion then involved George Airy as the Astronomer Royal, Warren de la Rue, John Herschel, William Huggins, Father Angelo Secchi, and others [13]. Much of the debate would once again transpire in *The Reader* [2]. In 1865, no less than ten letters appeared in the popular magazine and included contributions from Secchi himself [24–33]. Scientists took the controversy beyond conventional journals into the public forum.

With time, Dawes' view [13, 30] rose to prominence and the concept of "*willow leaves*" faded from solar physics. With respect to granulations, Dawes reminded his readers that: "*Their existence was well known to Sir W. Herschel*" [30]. He cited Herschel directly [30]: "*There is all over the Sun a great unevenness in the surface which has the appearance of a mixture of small points of an unequal light*" [34]. Dawes elaborated on his own position: "*I have proposed to term them granules or granulations, as more suitable than any more definite appellation, and therefore unlikely to mislead*" [30]. Nasmyth discovered nothing new [13, 18, 19], but he generated tremendous interest in the nature of solar granules. In turn, this prompted Secchi to put forth his solar model [16, 17]. Dawes did not live to see the resolution of the conflict.

As for Secchi, he observed both the granules and the intergranular lanes. He addressed the appearance of the solar surface as follows: "*The bottom of the solar disc appeared to be formed of a fine black mesh whose links were very thin and full of bright points. It was not so much the shape of the grid that surprised us — for we had seen it also at other times with older methods — as its blackness, which was truly extraordinary. It was such that we suspected some illusion, but in concentrating on certain darker points and finding them of unchanging and precise forms, we no longer remained in doubt about the reality of the aspect. Of this grid-like structure we can give an approximate idea in saying that the Sun looked like a ordinary piece of rough paper seen through a strong microscope; on this paper the prominences are numerous and irregular, and where the light falls rather obliquely, the bottom of the grooves are almost black compared to the more elevated parts, which appear extremely white... The grid-like solar structure seemed to us to offer nothing regular in those parts of the disc that are continuous, and thus the term granular appears very appropriate. The granular structure is more visible near the spots, but it is not recognizable in the faculae; these present themselves like luminous clusters without distinguishable separation, emitting continual light without the interruption of dots or of that black mesh. In the end, we have found the granular structure more notable and easy to distinguish in the middle of the disc than near the limb, and in the zones near the sun's equator, more than in the polar zones*" [17].

It was based on these observations that Secchi advanced his model of a gaseous Sun with a partially condensed photosphere: "*Indeed this appearance suggests to us what is perhaps a bold hypothesis. As in our atmosphere, when it is cooled to a certain point, there exists a fine substance capable of transforming itself in fine powder and of forming clouds in suspension, (water transforming into so-called "vesicular" vapor or into small solid icicles), so in the enflamed solar atmosphere there might be an abundance of matter capable of being transformed to a similar state at the highest temperatures. These corpuscles, in immense supply, would form an almost continuous layer of real clouds, suspended in the transparent atmosphere which envelopes the sun, and being comparable to solid bodies suspended in a gas, they might have a greater radiant force of calorific and luminous rays than the gas in which they are suspended. We may thus ex-*

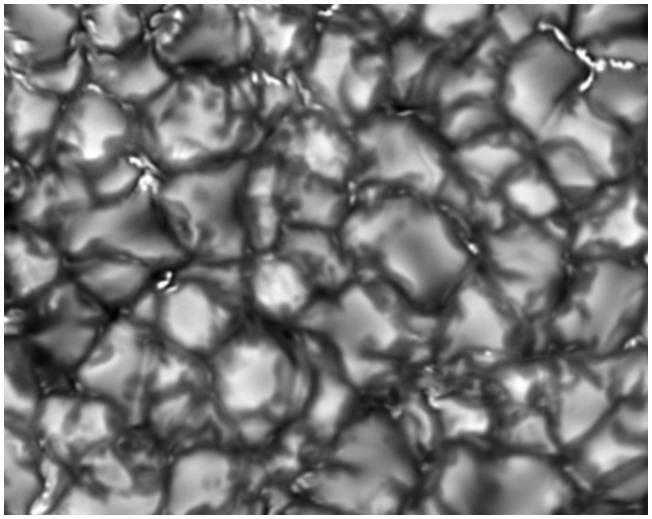


Fig. 1: High resolution image of solar granules acquired by Vasco Henriques on May 23, 2010 using the Swedish 1-m Solar Telescope (SST). "*The SST is operated on the island of La Palma by the Institute for Solar Physics of the Royal Swedish Academy of Sciences in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias*", http://www.solarphysics.kva.se.

plain why the spots (that are places where these clouds are torn) show less light and less heat, even if the temperature is the same. The excellent results obtained by Magnus, who has proved that a solid immersed in an incandescent gas becomes more radiant in heat and light than the same gas, seem to lend support to this hypothesis, which reconciles the rest of the known solar phenomena" [17]. With Secchi's words, others quickly followed suit [2, 25] and the Sun became viewed as having a gaseous body [2]. Such was the authority of Father Angelo Secchi in astronomy.

Objects which appeared as "*rice grains*" or "*willow leaves*" on the Sun's surface offered a rather poor foundation for scientific advancement. Chacornac would distance himself from these concepts: "*As to the form of the objects observed a subject so warmly discussed at the present time — I did not see, with the large instrument of the Paris Observatory, nor have I ever yet seen, that the form is limited to one only, either "willow leaf" or "rice grain". I have always seen the "crystals" of the photospheric atmosphere entangled (enchevêtrés) in a thousand ways, and connected among themselves by one or many points in their peripheries; I have always observed these photospheric clouds affecting forms reminding one of the flocculent mass in an incandescent metal, in suspension in a liquid... I have always in my descriptions compared the "crystals" of the photospheric matter to this silver solder in a state of fusion*" [25]. With these words, Chacornac became one of the first to invoke crystalline structure on the surface of the Sun. In the same letter [25], he echoed Secchi's model published in *Les Mondes* [17] three days prior, without properly referencing Secchi: "*... they constitute one of the essential conditions of the nature of this luminous matter, of which the elements are contained in the exterior atmosphere of the Sun as vapour is contained in our air*" [25]. Chacornac's description of the crystalline structure of granules would be revisited using theoretical analysis, more than 130 years later [35].

Scientists of the 19th century advocated that convection currents were the cause of granular formation. Gaseous material rose from deep within the Sun and then condensed on the photospheric surface before sinking once again in the gaseous atmosphere back towards the interior. The modern gaseous models promote similar hypotheses, but do not permit the condensation of matter. In 1881, Hastings described granules as follows: "*In our theory, then, the granules are those portions of upward currents where precipitation is most active, while the darker portions, between the bodies, are where the cooler products of this change with accompanying vapors are sinking to lower levels*" [36]. The convective nature of the granular field was well recognized, even though solar physicists lacked the mathematical tools required to address such problems.

2.2 Modern concepts of granules

The careful analysis of the solar granulation is important, as such studies reveal that the photosphere possesses objects with defined structures. The presence of such features provides compelling evidence that the Sun is constituted from condensed matter. Today, the study of solar granulation involves sophisticated image acquisition (see Figure 1) and data processing [14, 15, 35, 37–51]. Granules are widely regarded as the result of convective phenomena, wherein subsurface heat is being transported to the solar surface [14, 15, 37, 44, 50]. Convective processes move material upwards within the granule. Following radiative cooling, matter then sinks into the intergranule lanes [43]. The velocities of up and down flows can reach 1200 m/s in granular centers and intergranular lanes [43]. According to the gaseous models of the Sun, once the material reaches the surface layer, radiative heat losses result in greatly lowered opacity and the atmosphere of the Sun becomes transparent [37].

Granules vary in size from ~0.3–4 arcsec with most having a rough diameter of 1–2 arcsec giving a mean of ~1.35 arcsec (~1,000 km) [14, 37, 38]. Del Moro finds that no granule has an area larger than 1 Mm^2 [48]. Other investigators obtain maximal values in the 3-5 Mm^2 range [38, 45]. Small granules are very numerous, but they do not account for much of the solar surface [38]. They tend to be concentrated in downdraft regions, whereas the larger granules are located in areas of strong up currents [45]. The intergranular distance is on the order of 1.76 arcsec [38] and by some measures the darker intergranular lanes account for about 32% of the solar surface [42]. Conversely, Abdussamatov and Zlatopol'skii report that on a mesogranular scale (see below) the intergran-

ular lanes can occupy as much as 55% of the photospheric area [44]. Roudier and Muller provide an excellent review of many key facts relative to granules [38]. The structures tend to be irregular in shape, although they can be properly described as polygons with a slight prevalence of pentagons over hexagons [35].

If the log of the number of granules of a given size is plotted against the log of their area, two distinct lines can be used to fit all granules with a critical diameter of 1.31 arcsec (see Figure 7 in [38]). This suggest that "*granules are self-similar*" [15, 38, 45] which then implies structure. Smaller granules fit the first line and are thought to be produced by turbulent phenomena of a "*Kolmogorov-type*" [38]. Because they are believed to be the result of turbulent eddy motions, Roudier and Muller argue that these small structures should be viewed as "*photospheric turbulent elements*" [38], an idea consistent with their more prevalent occurrence in the downdraft regions [44]. Conversely, they state that only medium and larger structures should be viewed as true "*granules*" as these alone properly transport convective energy [38].

Mean granular lifetimes range from ~5 minutes to 16 minutes with a maximum of approximately 30 minutes [14, 46]. Granules are subject to three evolutionary mechanisms. Most often, they are produced through the fragmentation of larger systems [39, 40, 46, 49]. They often "*die*" through the merger of smaller entities [46]. They seldom appear from, but frequently dissolve into, the background [46]. The larger granules tend to have the largest lifetimes [48]. Granules that are "*long lived*" have a tendency to form clusters [49]. Dark dots often form within granules and these result in violent fragmentation of the structure producing "*exploding granules*" [39, 47, 51]. The formation of these dark dots results in fragmentation within a couple of minutes and the features have no link to magnetic fields [39,47]. Only very large granules explode [48]. Exploding granules are often very bright, initially suggesting the upward flow of matter followed by great expansion [39]. Their dark dots eventually evolve into intergranular dark regions which are indicative of downward flow even though some have argued, using opacity arguments, that dark dots represent upward material displacement [40].

Mark Rast proposed that exploding granules "*can be better understood if granulation is viewed as downflow-dominated-surface-driven convection rather than as a collection of more deeply driven upflowing thermal plumes*" [51]. Though not mentioned by Rast, such an idea would benefit from the presence of a real solar surface which only a condensed model of the Sun could provide [52].

The smaller the granule, the more likely it is to die without fragmentation or merging [40]. Conversely, if the granule is large, it is likely to merge or fragment [40]. The brightest region and the strongest upflows within large granules tend to be near the intergranular lanes and consequently are not located near the center of the structure [53]. A family of granules shares either fragmentation or merging and can have a lifetime approaching 46 minutes [40].

Granules can be organized into larger assemblies: mesogranulation, supergranulation, and giant cells [41–45]. Such assemblies share common and simultaneous changes in size, temperature, or other parameters [43]. Mesogranulation areas usually tend to be brighter, more dynamically active [42]. They are thought to represent a greater uplifting of matter and can span from 6–9 arcsec [43] and have lifetimes ranging from 30 minutes to 6 hours [48]. They are viewed as connected to common convective origins located at depths of 3,000–8,000 km [43]. Supergranular cells are believed to have their origins at depths of 20,000–30,000 km, while giant cells might stem from convective processes located as deep as 200,000 km below the surface [43]. These hypothetical depths are inherently linked to the gaseous models of the Sun.

Giant cells divide successively into supergranular and mesogranular structures [43]. However, Rast believes that mesogranulation and supergranulations are "*secondary manifestations of granulation itself*" [51]. He provides an excellent review of the solar granulation and these structures [53]. Granules tend to have limited vertical flows on the order of 1 km/s while the mesogranulation with their ~5,000–10,000 km diameters, can have flows approaching 60 km/s [53, 54]. Ikhsanov et al. suggest that the solar surface supports protogranules which are intermediate in size between granules and mesogranules [54]. Supergranulations possess diameters of ~30,000 km, display a 20 hour lifetime, and can manifest horizontal flows on the order of 400 km/s [53]. Such horizontal flows are contrary to a fully gaseous model of the Sun, as highlighted by the author (see §10 in [55]). Recently, Arkhypov et al. have found that Kolmogorov turbulence determines large scale surface activity on the photosphere [56] and claim these indicate that sub-surface convection motion can be detected through photospheric activity of supergiant complexes.

Granules display varying emissivities, but most studies simply report values for the granules and the intergranular lanes (e.g. [44] reports $+8\pm7.5\%$ for granules and $-7\pm5.5\%$ for the intergranular lanes). These descriptions appear to be over simplified as a smooth transition exists between the maximum brightness of a granule and the darkest point of the intergranular lane. As a result, considerable variability can be expected in such values.

Center to limb variations in granular intensities have also been investigated [57, 58]. Initially, Hidalgo et al. reported that granular contrast increased slightly towards the limb up $\mu = 0.6$, followed by a decrease in contrast moving further away from the solar center [57]. It is not clear if this change was due to an increase in brightness. Later, in a wavelength dependent study (0.8 μm and 1.55 μm), Cuberes et al. observed a monotonic decrease in contrast from the center of the solar surface ($\mu = 1$) towards the limb ($\mu = 0.3$) [58]. The change was steeper at the lower wavelength [58]. No peak was observed in contrast variation at either frequency [58]. The contrast at the center of the solar surface was dependent

on wavelength, with larger contrast (6.1%) at 0.8 μm, while only 2.9% at 1.55 μm [58].

Title et al. [59, 60] have studied the formation of granules in association with magnetic fields and discovered significant differences relative to size, intensity variation, and lifetime.

Recently, Getling et al. published a series of stunning reports implying that the solar surface possesses a series of ridges and trenches [61–64]. On first inspection, the results appear valid and the authors have gone through considerable lengths to eliminate artifacts [64]. If these findings are genuine, they suggest that the solar surface contains "*quaziregular*" structural systems of great breath and regularity [61–64]. Nonetheless, it is currently unclear if these fascinating results will withstand scrutiny. If so, they would constitute additional support for the condensed nature of the photosphere.

Solar granulations have been the subject of intense theoretical work (e.g. [65]). From the onset [66–68], such studies have been subject to the charge that they can, at times, constitute "*little more than an exercise in parameter fitting*" [67]. Clearly, the gaseous models of state do offer significant flexibility with respect to the number of usable parameters [69]. Given enough variables, fits can almost always be achieved. Nonetheless, this brief review of solar granulation reveals that these elements are filled with structural properties based on size, behavior, and lifetimes. In this regard, it is instructive to consider how solar granulations conform to the laws of convection, turbulence, and structure as obtained in condensed matter (see §3, §4, and §5).

3 Granules and the laboratory

The analysis of granulations as convective processes has always rested on the science of liquids. In 1900, Bénard convection was first observed in the liquid state [70, 71] and the process continues to be a property of condensed matter. Bray et al. re-emphasized that Bénard convection was dominated by surface tension, not buoyancy [14, p.116].

Bénard (or Bénard-Marangoni) convection [72–74] is characterized by hexagonal structures. In fact, such features are properties of both Bénard convection [70–74] and many solar granules [14]. It is difficult to discount the presence of these structural elements on the surface of the Sun as coincidental, even though many solar physicist deny the presence of Bénard convection. Yet, even the laws of Kolmogoroff turbulence are strictly applicable only to an *incompressible* fluid [14, p.14], a framework well-beyond that afforded by the gaseous Sun. Still, since solar physicists currently endorse a gaseous model of the Sun, granular convection has always been viewed as a buoyancy driven phenomenon. Bénard convection cannot occur on the surface of the Sun if a gaseous body is to be preserved. To propose otherwise automatically requires surface tension, an impossibility for gaseous models. Nonetheless, it is particularly troubling that most laboratory experiments used to treat granulation have been performed on incompressible liquids [14, p.116]. To avoid surface tension, experimentalists study incompressible liquids placed between rigid plates [14, p.116]. Such a setting is hardly the equivalent of the hypothetical and illusionary gaseous solar surface.

4 Granulations and crystal structure

Beyond these applications of liquids to the treatment of granular convection, Noever has used the methods of statistical crystallography to analyze the solar surface [35]. He has reported that the granular field displayed a remarkable similarity to crystals [35]. Solar granulation followed both the laws of Aboav-Weaire and of Lewis [75–77] for space filling structures in two dimensions. The agreement with the Aboav-Weaire law had an R value of 0.998, indicating "*a correlation which does not extend beyond the nearest neighbor cells*" [35]. Noever also found that granules followed the perimeter law, suggesting that many sided structures have larger perimeters (R = 0.987) [35]. Adherence to the perimeter law implied that "*energy is carried by the cell boundaries*" [35]. Noever stated: "*It is particularly noteworthy that prior to grain fragmentation, a dark region of low luminosity typically appears near this predicted low energy core of each cell. The perimeter law predicts this outcome derived not from any specific fluid parameters but from a statistical picture of lattices alone*" [35]. With these words, Noever accounted for the origin of exploding granules without any recourse to convection, based solely on structural energy considerations. Structure led to behavior and this directly implied that the granulations are condensed matter. Noever further demonstrated that granules obey Lewis' law which relates two dimensional area and cell sidedness (R = 0.984) [35]. This places a restriction on granulation based on the need to fill two dimensional space entirely [35]. Gases cannot assume two dimensional space filling forms and cannot follow the laws of structure. Liquids alone can truly account for the convective and structural nature of granules.

Regrettably, Noever's work has been largely neglected [78–81] receiving only one citation relative to solar science [81]. Nonetheless, it represented a critical contribution in the understanding of granulations, precisely because it implied that granules are condensed matter.

5 Emissivity: A common link for solar surface structures

5.1 Metals and sunspots

Non-metals are known to possess directional spectral emissivities which monotonically decrease with increasing angle as illustrated schematically in Fig. 2 [82–84]. Their normal emissivity is typically higher than their directional spectral emissivity. Conversely, metals tend to have lower normal spectral emissivities relative to their directional spectral emissivities. For metals, the directional spectral emissivities usually rise with increasing angle until they fall precipitously as

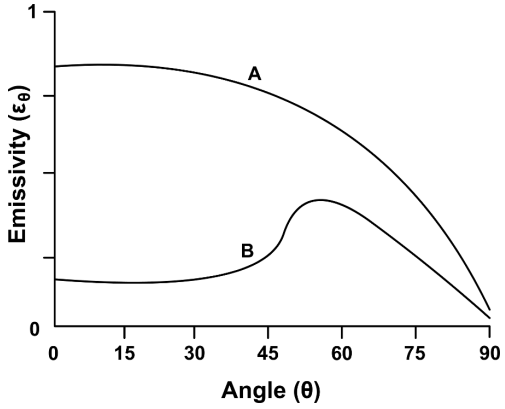


Fig. 2: Schematic representation of directional spectral emissivities for non-conductors (A) and conductors (B). Note that in non-metals, the spectral emissivity decreases monotonically with viewing angle. Conversely, in metals, while the normal emissivity can be substantially reduced, the emissivity can rise with increasing angle before precipitously dropping (adapted from [83]).

orthogonal viewing is approached [82,83]. These simple considerations provide tremendous insight to the structure of the photosphere in the context of a condensed solar model [52].

Consider the liquid metallic hydrogen model of the Sun [52]. When first proposed [85], liquid metallic hydrogen was hypothesized to assume a layered graphite-like structure. This lattice was subsequently adopted for the solar photosphere [52].

Since graphite itself is a great emitter, but only a modest conductor, one can hypothesize that liquid metallic hydrogen on the surface of the Sun is not highly metallic [52]. The inter-atomic distance in the lattice must be such that the photosphere displays little metallic character, but great graphite-like emissivity. This would correspond to the Type-I lattice structure previously discussed by the author [52, 55]. However, within sunspots, the interatomic distance would contract and liquid metallic hydrogen would increase its metallic character while at the same time, lowering its emissivity. In the limit, this would correspond to the Type II lattice [52].

The point can be amplified by examining the emissive behavior of sunspots with respect to magnetic field intensity [86, 87]. Leonard and Choudhary have reported that the emissivity of sunspot umbral regions drops with magnetic field strength suggesting the approach to a saturation limit (see Figure 2 in [86]). They stated: "*Although there is a large scatter, it is tempting to infer that the sunspot umbral intensity attains a maximum value beyond which the magnetic field increases without substantial intensity drop, resulting in a 'saturation effect'*" [86]. While more data of this nature is required, these preliminary findings imply that a limiting structural lattice might be reached within sunspots.

Sunspots are known to have substructure [88] and, as they can be the source of powerful magnetic fields [89], such observations [86] further support the notion that they are metallic in character [52]. The dark nuclei of sunspots clearly have lower emissivities and possess the highest magnetic fields [8, p.80]. Conversely, the light bridges display higher emissivities and lower magnetic fields [8, p.85–86], implying that they are less-metallic in character. The dark cores detected in sunspot penumbral filaments might be a reflection of increased metallicity in these elements [90].

Supportively, helioseismology reveals that sound waves travel much faster through sunspots than through normal photospheric matter [91, 92]. This suggests that the modulus of elasticity is higher within sunspots, in accordance with the hypothesis that the material is both more metallic and slightly denser than photospheric matter.

Consequently, greater attention might be placed on evaluating directional emissivity within sunspots. Measurements from these regions are already giving hints that emissivities may be increasing with angle of visualization. This is reflected in the "*problems of stray light*" into the sunspots [8, p.75–77]. The effect of "*stray light*" acts to increase the observed emissivity of sunspots in precisely the same manner that an increased metallic character would produce (see Eq. 8 in [8, p.75]). As a result, such data may already be affirming the metallic character of sunspots by mimicking the behavior manifested in Fig. 2. "*Stray light*" arguments might have been introduced simply to address a finding which could not be explained otherwise by the gaseous solar models. The observation of large sunspots at high resolution should enable scientists to clearly establish the directional emissivity of sunspots without any "*stray light*" effects and thereby possibly affirm their metallic nature.

It is appropriate to consider that sunspots might represent liquid metallic hydrogen whose lattice density has increased along with a corresponding rise in metallic nature: the stronger the metallic character, the stronger the associated magnetic fields and the weaker the emitted light intensity. This is precisely what one observes in sunspots [86]. Emissivity is strongly dependent on magnetic field intensity. As magnetic field intensity increases, sunspot emissivity progressively falls until a plateau region appears to be reached [86]. This would correspond to the limit of compressibility of the lattice. Beyond this point, liquid metallic hydrogen should become essentially incompressible, the Type II lattice having been reached.

Along these lines, it is interesting to note that liquid graphite displays two lattice forms which differ in spatial dimensions, densities, and metallic character [93]. Liquid graphite [93] appears to provide an interesting parallel with the two structural lattice Types required in a liquid metallic hydrogen model of the Sun [52].

These results can only be explained with difficulty using the gaseous models. After all, the presence of magnetic fields by themselves can have no effect on emissivity. It is well known that a piece of iron does not change its emissivity on

becoming magnetized. Emissivity changes demand changes in structure [94] and the gaseous solar models afford none.

5.2 Emissivity, granulation, and limb darkening

Frank Very was the first to monitor the limb darkening of the Sun [12] as a function of frequency. Very examined the photosphere at 7 wavelengths ranging from 0.416 μm to 1.5 μm [12]. He found that limb darkening was much more pronounced at shorter wavelength [12]. Since that time, extensive studies of limb darkening have been performed (e.g. [9–11]). Pierce and his collaborators provided an detailed list of coefficients for polynomial representations of limb darkening spanning a wide range of frequencies [9, 10]. Overall, these functions demonstrated that the photosphere behaves as a non-metal.

Today, limb darkening constitutes a central pillar of the gaseous solar models. The phenomenon remains linked to solar opacity arguments [95]. Nonetheless, when Very first considered the frequency dependence of limb darkening [12], he did not ponder only upon opacity arguments. He questioned whether limb darkening could be explained by the granulated aspect of the solar surface [12].

Solar granules display emissive characteristics which change towards their periphery as the dark intergranule lanes are reached. They also display center to limb variations [57, 58]. In fact, it is likely that the same phenomenon is being observed both locally near the granules and over the expanse of the entire solar surface as the limb is visualized. Granules manifest a brightness which fades in the intergranule lanes in the same manner as darkening manifests itself from the center of the solar body to the limb. As such, higher spatial resolution on granules may soon reveal that they individually exhibit the same features as observed globally in limb darkening. This would be expected if the emissivity of the Sun simply reflected the constitution of its condensed surface. Each individual granule would become a local manifestation of the limb darkening observed over the entire solar disk.

6 Conclusions

From the days of their discovery by William Herschel [34], granules have offered solar science a vast and fascinating array of structural forms which follow specific evolutionary paths and predetermined timelines. By every measure, granules are real entities, not illusions. They obey the laws of two-dimensional structures and manifest themselves as objects which can be analyzed, categorized, and mathematically evaluated. They appear and behave as condensed matter.

Conversely, a gaseous Sun should be devoid of structural elements: sunspots, granules, prominences, and flares which rupture the solar surface. It should be a blob, a haze, a nondescript mass — not a body filled with structure, as Secchi so elegantly described in his classic text [21]. A brief study of granulations and sunspots demonstrates that these are real structures which follow in every manner the behavior of condensed matter. The issues are not only structural, but involve the ability to have variable emissivities and powerful magnetic fields. On the Earth, the generation of strong magnetic fields remains associated with metallic character [55]. Gases can never produce magnetic fields of themselves. They simply respond to such phenomena.

The fact that sunspots possess strong magnetic fields might guide the synthesis of liquid metallic hydrogen on the Earth [52]. If the Sun is really made of liquid metallic hydrogen, then our study of sunspots implies that the material is easily endowed with magnetic properties. Therefore, it is possible that the synthesis of metallic hydrogen on the Earth could benefit by placing the entire experimental setting within a modest magnetic field on the order of 0.5 Tesla. This would correspond to the maximal 5,000–6,000 gauss field observed within sunspots [86, 87]. Large bore human magnetic resonance imaging (MRI) magnets currently operate up to fields of 9.4 Tesla, thereby confirming that suitable magnet technology exists for such studies [96].

At the same time, it is clear that the proper study of granular and sunspot emissivity will require much stronger optical space telescopes devoid of the "*seeing problems*" [1, p.23–25] when visualizing the Sun from the Earth. Resolutions must be increased tremendously such that emissivity can be properly mapped across an individual granule or sunspot umbra. When studying granulations, such maps should be married with Doppler imaging of the solar topology in order to link emissivity to angular changes in the surface. In this manner, solar physicists should be able to directly associate observed darkening with the emissive behavior of the solar surface itself, whether locally on the granular scale, or globally, as observers compare the solar center to the limb. In addition, the study of directional emissivities in sunspots should eventually affirm their metallic nature making investments in powerful space solar telescopes vital to the proper understanding of the solar surface.

As we continue to ponder the nature of the Sun, it is appropriate to close by recalling the brilliance of Father Secchi as an astronomer. Above all, Secchi valued observations. He painstakingly generated drawings of the Sun in an attempt to describe solar structures. Through his writings, he demonstrated that observation must lead theory. Short of data, we know nothing of the Sun. Therefore, should solar physics advance, the tradition of careful observation which Secchi inspired must be imitated. Even 140 years after the publication of *Le Soleil* [21], Secchi continues to astound, as Sobotka highlights [8]: "*In 1870 appeared the first edition of a fundamental work in solar astronomy by P. A. Secchi: Le Soleil. Most of the basic concepts of the sunspots' morphology can be found there. Secchi made his visual observations from 1865 to 1870 with a resolution approaching to 0″.3 in some cases. In his wonderful drawings he presented not only the basic morphological features like multiple umbrae,*

light bridges, and penumbral filamentary structure, but also "knots" in bright penumbral filaments (penumbral grains) and internal structure of light bridges. He also noticed spatial variations in umbral brightness and the darkest regions — "holes" — in the umbra (dark nuclei). In three of his drawings even some umbral dots can be seen, although he did not describe them". Now, endowed with the gifts of modern technology, solar physicists must be better equipped to properly describe what Secchi himself could only observe in awe using a simple telescope.

Acknowledgement

The author would like to thank Luc Robitaille for the preparation of Figure 2.

Dedication

Cet ouvrage est dédié à celles qui ont été parmi mes premières enseignantes, les Filles de la Sagesse.

Submitted on June 24, 2011 / Accepted on July 03, 2011
First published online on July 16, 2011

References

1. Vaquero J.M. and Vázquez M. The Sun recorded through history: scientific data extracted from historical documents. Astrophysics & Space Science Library, v.361, Springer, Dordrecht, 2009.
2. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v.3, 3–25 — A paper published in this Special Issue.
3. Galilei G. and Scheiner C. On Sunspots. Translated by E. Reeves and A.V. Helden, University of Chicago Press, Chicago, 2010.
4. Vaquero J.M. Historical sunspot observations: A review. *Adv. Space Res.*, 2007, v.40(7), 929–941.
5. Maunder E.W. The Sun and sunspots. *Mon. Not. Roy. Astron. Soc.*, 1922, v.82(9), 534–543.
6. Bray R.J. and Loughhead R.E. Sunspots: The international astrophysics series, volume seven. Chapman and Hall Ltd, London, 1964.
7. Robitaille P.M. A thermodynamic history of the solar constitution — II: The theory of a gaseous Sun and Jeans' failed liquid alternative. Progr. Phys., 2011, v.3, 41–59 — A paper published in this Special Issue.
8. Sobotka M. Fine structure in sunspots. In: *Motions in the solar atmosphere* (A. Hanslmeier and M. Messerotti, eds.), Astrophysics and Space Science Library, v.239, Kluwer Academic Publishers, Dordrecht, 1999, p.71–97.
9. Pierce A.K. and Slaughter C.D. Solar limb darkening I. ($\lambda\lambda$3033 to 7297). *Solar Physics*, 1977, v.51, 25–41.
10. Pierce A.K., Slaughter C.D., and Weinberger D. Solar limb darkening in the interval 7404-24 018 Å. II. *Solar Physics*, 1977, v.52, 179–189.
11. Neckel H. and Labs D. Solar limb darkening 1986–1990 ($\lambda\lambda$303 to 1099 nm). *Solar Physics*, 1994, v.153(1–2), 91–114.
12. Very F.W. The absorptive power of the solar atmosphere. *Astrophys. J.*, 1902, v.16, 73–91.
13. Bartholomew C.F. The discovery of the solar granulation. *Quart. J. Roy. Astron. Soc.*, 1976, v.17, 263–289.
14. Bray R.J., Loughhead R.E., Durrant C.J. The solar granulation. Cambridge University Press, Cambridge, 1984.
15. Muller R. The solar granulation. In: *Motions in the solar atmosphere* (A. Hanslmeier and M. Messerotti, eds.), Astrophysics and Space Science Library, v.239, Kluwer Academic Publishers Dordrecht, 1999, p.35–70.
16. Secchi A. Sulla Teoria Delle Macchie Solari: Proposta dal sig. Kirchoff. *Bullettino Meteorologico dell'Osservatorio del Collegio Romano*, 31 January 1864, v.3(4), 1–4 (translated into English by Eileen Reeves and Mary Posani: On the theory of sunspots proposed by Signor Kirchoff. *Progr. Phys.*, 2011, v.3, 26–29 — A paper published in this Special Issue).
17. Secchi A. Sulla Struttura della Fotosfera Solare. *Bullettino Meteorologico dell'Osservatorio del Collegio Romano*, 30 November 1864, v.3(11), 1–3 (translated into French by François Moigno: Sur la structure de la photosphère du soleil. *Les Mondes*, 22 December 1864, v.6, 703–707; translated into English by Mary Posani and Eileen Reeves: On the structure of the photosphere of the Sun. *Progr. Phys.*, 2011, v.3, 30–32 — A paper published in this Special Issue).
18. Nasmyth J. On the structure of the luminous envelope of the Sun — In a letter to Joseph Sidebotham. *Mem. Lit. Phil. Soc. Manchester*, 1862, 3rd Series, v.1, 407–411.
19. Nasmyth J. The willow leaves. *Astron. Reg.*, 1865, v.3, 223–224.
20. Magnus G. Notice sur la constitution du soleil. *Archives des science physique et naturelles* (Genève), 1864, v.20, 171–175; (translated from: Magnus G. Notiz ber die Beschaffenheit der Sonne. *Poggendorff's Annalen der Physik und Chemie*, 1864, v.131, 510–512; translated into English by Patrice Robitaille: Notice on the constitution of the Sun. *Progr. Phys.*, 2011, v.3, 33–34 — A paper published in this Special Issue).
21. Secchi A. Le Soleil, Gauthier-Villars, Paris, 1870.
22. Smiles S. James Nasmyth: An autobiography. John Murray, London, 1883.
23. Denning W.F. The Rev. William Rutter Dawes. *The Observatory*, 1913, v.36(467), 419–422.
24. Unsigned. Recent astronomical literature. *The Reader*, 1865, v.5, 8–9.
25. Chacornac, On the physical constitution of the Sun. (25/12/1864). *The Reader*, 1865, v.5, 16–17 (Note: This letter was drafted from Lyon just three days after Secchi's classic paper was published in *Les Mondes* [2, 12]).
26. Secchi A. Father Secchi on "Willow Leaves". (22/2/1865). *The Reader*, 1865, v.5, 259.
27. Secchi A. On the solar surface. (6/7/1865). *The Reader*, 1865, v.6, 125.
28. Unsigned. Solar observations. *The Reader*. 1865, v.6, 152–153.
29. Airy G.B. On the solar surface. (2/8/1865). *The Reader*, 1865, v.6, 155.
30. Dawes W.R. The solar surface. (8/16/1865). *The Reader*, 1865, v.6, 210–211.
31. Stone E.J. Untitled. *The Reader*, 1865, v.6, 211.
32. Secchi A. Untitled. (8/8/1865). *The Reader*, 1865, v.6, 211.
33. Secchi A. The solar surface. (11/8/1865). *The Reader*, 1865, v.6, 241–242.
34. Herschel W. On the nature and construction of the Sun and fixed stars. *Phil. Trans. Roy. Soc. London*, 1795, v.85, 46–72.
35. Noever D.A. Solar granulation and statistical crystallography: a modeling approach using size-shape relations. *Astron. Astrophys.*, 1994, v.282, 252–261.
36. Hastings C.S. A theory of the constitution of the Sun, founded upon spectroscopic observations, original and other. *Am. J. Science*, 1881, v.21, 121–132.
37. Hirzberger J. Granulation and waves. *Astron. Nachr.*, 2003, v.324(4), 344–348.

38. Roudier Th. and Muller R. Structure of the solar granulation. *Solar Physics*, 1986, v.107, 11–26.
39. Kitai R. and Kawaguchi I. Morphological study of the solar granulation I: Dark dot formation in the cell. *Solar Physics*, 1979, v.64, 3–12.
40. Kawaguchi I. Morphological study of the solar granulation II: The fragmentation of granules. *Solar Physics*, 1980, v.65, 207–220.
41. November L.J., Toomre J., Gebbie K.B. and Simon G.W. The detection of mesogranulation on the Sun. *Astrophys. J.*, 1981, v.245, L123–L126.
42. Oda N. Morphological study of the solar granulation III. The mesogranulation. *Solar Physics*, 1984, v.93, 243–255.
43. Abdussamatov H.I. The fine structure of solar granulation and its relationship to large-scale photospheric structure. *Astron. Astrophys.*, 1993, v.272, 580–586.
44. Abdussamatov H.I. and Zlatopol'skii A.G. Variation of fine-structure properties of solar granulation on a mesogranular scale. *Astron. Letters*, 1997, v.23(6), 752–757.
45. Hirzberger J., Vázquez M., Bonet J.A., Hanslmeier A. and Sobotka M. Time series of solar granulation images I: Differences between small and large granules in quiet regions. *Astrophys. J.*, 1997, v.480, 406–419.
46. Hirzberger J., Bonet J.A., Vázquez M., and Hanslmeier A. Time series of solar granulation images II: Evolution of individual granules. *Astrophys. J.*, 1999, v.515, 441–454.
47. Hirzberger J., Bonet J.A., Vázquez M. and Hanslmeier A. Time series of solar granulation images III: Dynamics of exploding granules and related phenomena. *Astrophys. J.*, 1999, v.527, 405–414.
48. Del Moro D. Solar granulation properties derived from three different time series. *Astron. Astrophys.*, 2004, v.428, 1007–1015.
49. Dialetis D., Macris C., Muller R. and Prokakis T. A possible relation between lifetime and location of solar granules. *Astron. Astrophys.*, 1988, v.204, 275–278.
50. Hirzberger J. On the brightness and velocity structure of solar granulation. *Astron. Astrophys.*, 2002, v.392, 1105–1118.
51. Rast M.P. On the nature of "exploding" granules and granule fragmentation. *Astrophys. J.*, 1995, v.443, 863–868.
52. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v.3, 60–74 — A paper published in this Special Issue.
53. Rast M.P. The scales of granulation, mesogranulation, and supergranulations. *Astrophys. J.*, 2003, v.597, 1200–1210.
54. Ikhsanov R.N., Parfinenko L.D. And Efremov V.I. On the organization of fine structure of the solar photosphere. *Solar Physics*, 1997, v.170, 205–215.
55. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v.2, 17–21.
56. Arkhypov O.V., Antonov O.V. and Khodachenko M.L. Supergiant complexes of solar activity and convection zone. *Solar Physics*, 2011, v.270, 1–8.
57. Rodriguez-Hidalgo I., Collados M. and Vázquez M. Center-to-limb variation of solar granulation along the equator and the cental meridian. *Astron. Astrophys.*, 1992, v.264, 661–672.
58. Cuberes M.S., Vasquez M., Bonet J.A., and Sobotka M. Center-to-limb variation of solar granulation in the infrared. *Astron. Astrophys.*, 2003, v.397, 1075–1081.
59. Title A.M., Tarbell T.D., Simon G.W. Acton L., Duncan D., Ferguson S., Finch M., Frank Z., Kelly G., Lindgren R., Morrill M., Pope T., Reeves R., Rehse R., Shine R., Topka K., Harvey J., Leibacher J., Livingston W., November L. White-light movies of the solar photosphere from the soup instrument on Spacelab 2. *Adv. Space Res.*, 1986, v.6(8), 253–262.
60. Tittle A.M., Tarbell T.D., Topka K.P., Ferguson S.H., Shine R.A. and the SOUP Team. Statistical properties of solar granulation derived from the SOUP instrument on Spacelab 2. *Astrophys. J.*, 1989, v.336, 475–494.
61. Getling A.V. and Brandt P.N. Regular structures of the solar photosphere. (Persistence of the granular field and trenching in the brightness relief.) *Astron. Astrophys.*, 2002, v.382, L5–L8.
62. Getling A.V. Do-quazi regular structures really exist in the solar photosphere I. Observational Evidence. *Solar Physics*, 2006, v.239, 93–111.
63. Getling A.V. Widespread occurrence of trenching patterns in the granulation field: Evidence for roll convection? *Solar Physics*, 2008, v.248, 233–245.
64. Getling A.V. Do long-lived features really exist in the solar photosphere? II. Contrast of time-averaged granulation images. *Solar Physics*, 2008, v.249(2), 307–314.
65. Stein R.F. and Nordlund A. Simulations of solar granulation I. General properties. *Astrophys. J.*, 1998, v.499, 914–933.
66. Margrave T.E. and Swihart T.L. Inhomogeneities in the Solar Photosphere. *Solar Physics*, 1969, v.6, 12–17.
67. Wilson P.R. On granulation models. *Solar Physics*, 1969, v.8, 20–22.
68. Margrave T.E. and Swihart T.L. More on granulation models. *Solar Physics*, 1969, v.9, 315–316.
69. Nelson G.D. A two-dimensional solar model. *Solar Physics*, 1978, v.60, 5–18.
70. Bénard H. Les tourbillons céllulaires dans une nappe liquide. *Rev. Gen. Sci. Pures Appl.*, 1900, v.11, 1261–1271; 1309–1328.
71. Bénard H. Les toubillons céllulaires dans une nappe liquid transportant de la chaleur par convection en régime permanent. *Ann. Chim. Phys.*, 1901, v.23, 62–144.
72. Ozbelge H.O. On the surface structure and the hydrodynamics of the Bénard Cells. *Experiment. Fluids*, 1989, v.8, 238–240.
73. Cerisier P., Rahal S., Rivier N. Topological correlations in Bénard-Marangoni convective structures. *Phys. Rev. E.*, 1996, v.54(5), 5086–5094.
74. Thiele U. and Eckert K. Stochastic geometry of polygonal networks an alternative approach to the hexagon-square-transition in Bénard convection. *Phys. Rev. E*, 1998, v.58, 3458–3468.
75. Weaire D. and Rivier N. Soap, cells and statistics random patterns in two dimensions. *Contemp. Phys.*, 1984, v.25, 59–99.
76. Rivier N. Structure of random cellular networks. Science on Form: *Proc. First Int. Symp. Sci. Form*, KTK Scientific Publishers, Tokyo, 1986, 451–458.
77. Chiu S.N. Aboav-Weaire's and Lewis' Law A review. *Mater. Charact.*, 1995, v.34, 149–165.
78. Bandeira L., Pina P. and Saraiva J. A multi-layer approach for the analysis of neighbourhood relations of polygons in remotely acquired images. *Pattern Recogn. Letters*, 2010, v.31(10), 1175–1183.
79. Saraiva J., Pina P., Bandeira L. and Antunes J. Polygonal networks on the surface of Mars; applicability of Lewis, Desch and Aboav-Weaire laws. *Phil. Mag. Letters*, 2009, v.89(3), 185–193.
80. Pina P, Saraiva J, Bandeira L. and Antunes J. Polygonal terrains on Mars: A contribution to their geometric and topological characterization. *Planetary Space Sci.*, 2008, v.56(15), 1919–1924.
81. Berrilli F, Florio A, Consolini G, Bavassano B., Briand C., Bruno R., Caccin B., Carbone V., Ceppatelli G., Egidi A., Ermolli I., Mainella G. and Pietropaolo E. Dependence of the photospheric vertical flow characteristics on the granule dimension. *Astron. Astrophys.*, 1999, v.344(2), L29–L32.
82. Modest M.F. Radiative heat transfer. McGraw-Hill, New York, 1993, p.92–108.

83. Thirumaleshwar M. Fundamentals of Heat and Mass Transfer, Dorling Kindersley, Dehli, 2009, p.652.
84. Incropera F.P., DeWitt D.P., Bergman T.L., Lavine A.S. Fundamentals of Heat and Mass Transfer, 6th Edition. John Wiley & Sons, Hoboken, NJ, 2007.
85. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v.3, 764–770.
86. Leonard T. and Choudhary D.P. Intensity and magnetic field distribution of sunspots. *Solar Physics*, 2008, v.252, 33–41.
87. Martinez Pillet V. and Vázquez M. The continuum intensity-magnetic field relation in sunspot umbrae. *Astron. Astrophys.*, 1993, v.270, 494–508.
88. Moradi H., Baldner C., Birch A.C., Braun D.C., Cameron R.H., Duvall T.L., Gizon L., Haber D., Hanasoge S.M., Hindman B.W., Jackiewicz J., Khomenko E., Komm R., Rajaguru P., Rempel M., Roth M., Schlichenmaier R., Schunker H., Spruit H.C., Strassmeier K.G., Thompson M.J. and Zharkov S. Modeling the subsurface structure of sunspots. *Solar Physics*, 2010, v.267, 1–62.
89. Livingston W., Harvey J.W., Malanushenko O.V. and Webster L. Sunspots with the strongest magnetic fields. *Solar Physics*, 2006, v.239, 41–68.
90. Scharmer G. B., Gudiksen B.V., Kiselman D., Lfdahl M. G., Rouppe van der Voort L. H. M. Dark cores in sunspot penumbral filaments. *Nature*, 2002, v.420, 151–153.
91. Moradi H. and Cally P.S. Time-distance modeling in a simulated sunspot. *Solar Physics*, 2008, v.251, 309–327.
92. Ilonidis S. and Zhao J. Determining absorption, emissivity reduction, and local suppression coefficients inside sunspots. *Solar Physics*, 2011, v.268, 377–388.
93. Zazula J.M. On graphite transformations at high temperature and pressure induced by absorption of LHC beam. LHC Project Note 78/97 (January 18, 1997). available online: cdsweb.cern.ch/record/691793/files/project-note-78.pdf.
94. Touloukian Y.S. and Ho C.Y. Thermophysical Properties of Matter (vols. 1). Plenum, New York, 1970.
95. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v.3 — A paper published in this Special Issue.
96. Robitaille P.M. and Berliner L.J. Ultra high field magnetic resonance imaging. Springer, New York, 2006.

On the Temperature of the Photosphere: Energy Partition in the Sun

Pierre-Marie Robitaille
Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA
E-mail: robitaille.1@osu.edu

In this note, energy partition within the Sun is briefly addressed. It is argued that the laws of thermal emission cannot be directly applied to the Sun, as the continuous solar spectrum ($T_{app} \sim 6,000\,\mathrm{K}$) reveals but a small fraction of the true solar energy profile. Without considering the energy linked to fusion itself, it is hypothesized that most of the photospheric energy remains trapped in the Sun's translational degrees of freedom and associated convection currents. The Sun is known to support both convective granules and differential rotation on its surface. The emission of X-rays in association with eruptive flares and the elevated temperatures of the corona might provide some measure of these energies. At the same time, it is expected that a fraction of the solar energy remains tied to the filling of conduction bands by electrons especially within sunspots. This constitutes a degree of freedom whose importance cannot be easily assessed. The discussion highlights how little is truly understood about energy partition in the Sun.

The discussion of energy partition in materials may be considered to be so complex at times that, perhaps, the most prudent course of action rests in avoiding the entire subject. In the laboratory, the evaluation of energy partition demands years of study involving many hurdles for meager rewards. Nonetheless, before progress can be made in any field, the issues at hand must be identified. It is worthwhile to highlight some general ideas relative to energy partition in the Sun which would eventually afford a detailed mathematical approach to the question. Relative to solar physics, energy partition is complicated by the presence of both conduction and convection on the solar surface.

The interior of the Sun is currently hypothesized to approach temperatures of ~15,600,000K, while the corona manifests values on the order of 2,000,000–3,000,000 K [1, p.10]. Solar physicists maintain that the solar photosphere exists at a temperature of ~5,780K [1, p.10] in an apparent violation of the second law of thermodynamics [2–4]. This surface temperature is based on the application of the laws of thermal emission [5–7] to the solar spectrum [1, p.3–9] as first recorded in its entirety by Langley [8–10]. Still, the assignment of a temperature to the photosphere has not been without controversy.

Throughout the 19th century, great variations existed with respect to the temperature of the photosphere (see [11, p.268–279] and [12, p.48–52] for reviews). In 1898, Scheiner brought apparent unification to the problem when he applied Stefan's law [6] to data acquired by Pouillet, Secchi, Violle, Soret, Langley, Wilson, Gray, Paschen, and Rosetti [13]. Scheiner demonstrated that these previously discordant studies (see [14] for many of the original values) resulted in calculated solar temperatures of 5,000 to 6,200 K, with only one observation standing at 10,000 K [13]. Scheiner believed in a gaseous model and insisted that, even though the Sun's layers supported differing temperatures, it might be viewed as a blackbody. However, such an object did not meet the equilibrium conditions required by Kirchhoff [15, 16]. This immediately brought into question any temperature derived from such methods.

Scheiner was not alone in advocating that the laws of thermal radiation could be applied to the Sun. Two years earlier, in order to justify the extraction of the photospheric temperature from the laws of thermal radiation, Ebert stated that: "*With respect to electromagnetic radiation, the principal mass of the Sun acts like a black body*" [17]. In 1895, most scientists believed that Secchi's model of the Sun [18,19] was valid. Ebert considered this framework when he initially expressed doubt about the blackbody nature of the Sun: "*There remains only the question, whether we can regard the incandescent particles of the Sun, which yield the continuous spectrum, as comparable to a black body with respect to their total radiating capacity*" [17]. Frank Very [20] was more adamant in questioning the applications of the laws of emission to solar data when, in 1908, he stated in *Science*: "*It is doubtful whether radiation formulae obtained from measures through a limited range of temperature for solid bodies, composed of complex molecules, are applicable to solar conditions at the photospheric level, where it is improbable that any molecules remain undissociated. Extrapolations from Stefan's law of the proportionality of total radiation from a black body to the fourth power of the absolute temperature, are therefore not certainly applicable to the problem, even though the law has been verified through a range of some hundreds of degrees*" [20]. Nonetheless, Very immediately applied Stefan's law to the Sun [20].

The sternest warning against applying the laws of radiation to the Sun would come from Max Planck [21]. The father of modern physics removed all doubt relative to his position when he wrote: "*Now the apparent temperature of the Sun is obviously nothing but the temperature of the solar rays, de-*

pending entirely on the nature of the rays, and hence a property of the rays and not a property of the Sun itself. Therefore it would be not only more convenient, but also more correct, to apply this notation directly, instead of speaking of a fictitious temperature of the Sun, which can be made to have meaning only by the introduction of an assumption that does not hold in reality" [22, §101]. If Planck was so forceful in his comment, he rested his case on solid grounds: "*It is only in the case of stable thermodynamic equilibrium that there is but one temperature, which then is common to the medium itself and to all rays whatever color crossing it in different directions*" [22, §101]. Planck recognized with these words that the Sun was not in thermal equilibrium and hence he refused to accept the concept of "*apparent*" or "*effective*" solar temperatures [22, §101].

Perhaps more than anyone, Max Planck recognized that the laws of thermal emission had been obtained in settings involving complete thermal equilibrium. Kirchhoff's formulation was restricted to radiation within a rigid enclosure [15, 16,22] sustaining full thermal equilibrium. There could be no net conduction or convection processes present. Based on his objection, Planck recognized that the Sun supported convection currents. Carrington's differential solar rotation had been well known for over fifty years [18] and the convective nature of granular field was also firmly established [23]. In view of Planck's warning, a more considered approach should be adopted relative to applying the laws of thermal emission to the Sun.

Max Planck specifically excluded conduction when treating radiation, on the grounds that it's presence violated thermal equilibrium: "*Now the condition of thermodynamic equilibrium requires that the temperature shall be everywhere the same and shall not vary with time. Therefore in any given arbitrary time just as much radiant heat must be absorbed as is emitted in each volume-element of the medium. For the heat of the body depends only on the heat radiation, since, on account of the uniformity in temperature, no conduction of heat can take place*" [22, §25]. Like conduction, convection reduces emissivity. It is known that the emissivity of gases can fall with temperature in clear violation of Stefan's law [24]. These two realities, the presence of conduction and convection on the photosphere, are likely to explain Planck's hesitation to state anything about the Sun, based solely on the acquisition of its spectrum. Nonetheless, perhaps it is possible to extract something of value from the solar spectrum with respect to energy distribution within the Sun.

Relative to thermal radiation, the availability of electrically conductive paths can alter emissivity. In metals, normal emissivity can be substantially reduced [25–27]. Silver is an excellent conductor, but a poor emitter [28]. In fact, polished silver has one of the highest coefficients of reflection. It can be concluded that electronic conduction reduces emissivity.

When energy enters or escapes from an object, it does so by filling or vacating available degrees of freedom [24]. Without considering nuclear processes, the degrees of freedom are either translational, vibrational, rotational, or electronic [24]. As a rule, electronic degrees of freedom become particularly important at elevated temperatures. Within a gaseous Sun, constituent atoms are viewed as existing in a dissociated state. Such monoatomic species can have recourse only to translational and electronic degrees of freedom. Vibrational and rotational degrees of freedom are restricted to species which are at least diatomic.

In a solid, such as graphite at room temperature, the dominant degrees of freedom are likely to be vibrational [24]. Graphite displays a reasonable thermal conductivity in the hexagonal plane (390 W/m×K for *ab* direction) [29, p.44–57]. This compares well with the thermal conductivity of silver (420 W/m×K) [29, p.57]. Conversely, the thermal conductivity of graphite drops substantially between layers (∼2 W/m×K) [29, p.57]. In graphite, thermal conductivity is linked to the vibrations of the lattice and these degrees of freedom [29, p.56].

Relative to electrical conductivity, graphite is a "semimetal" [29, p.57]. Its resistivity is $\sim 3{\times}10^{-3}$ ohm×m between layers making it is good insulator [29, p.61]. However, in the hexagonal plane, graphite has a resistivity of approximately $2.5{-}5{\times}10^{-6}$ ohm×m [29, p.61] making it reasonably metallic, but still well below silver which has an electrical resistivity of $\sim 1.59{\times}10^{-8}$ at 293 K [30, p.12–40]. Even in its favored plane, graphite is a significantly inferior conductor relative to silver. Consequently, the electrical conductivity of silver must be responsible for its weak emissivity, since its thermal conductivity is similar to graphite at least in one plane. This leads to the conclusion that the vibrational degrees of freedom are responsible for the excellent emissivity of graphite. Assuming that the object is at rest, the graphitic lattice does not permit translations or rotations, while the electronic degrees of freedom are unlikely to be significantly populated. As a result, when emissivity is properly coupled to temperature, it appears that the vibrational state of the sample primarily dominates [24].

In the gaseous models of the Sun, hydrogen and helium must exist as isolated atoms, many of which are devoid of electrons. Since the gaseous Sun has no lattice, it cannot support either thermal conduction through such a structure or energy transfer through electronic conduction bands. It cannot have recourse to lattice vibrations as a degree of freedom. Consequently, a gaseous Sun must rely almost exclusively on translational and electronic degrees of freedom as receptacles for energy. Yet, laboratory experience dictates that these degrees of freedom cannot support thermal emission of a Planckian nature [7]. Such is the great flaw of gaseous models which solar opacity approaches cannot reconcile [31]. To explain solar thermal emission, a mechanism similar to that which exists in graphite must be invoked. The dominant degrees of freedom in graphite are vibrational and linked to the existence of the lattice itself. In contrast, a gaseous Sun has

no lattice and therefore cannot produce a thermal spectrum. Opacity arguments do not suffice to rectify these problems in a gaseous solar model [31].

Conversely, within a liquid metallic hydrogen model of the Sun [32], a lattice exists. In fact, from the days when it was first proposed by Wigner and Huntington [33], metallic hydrogen has been hypothesized to be able to assume a layered lattice similar to graphite. Such a lattice configuration will possess vibrational degrees of freedom which mimic those found in graphite, as required to properly account for the production of the solar spectrum. Accordingly, the thermal spectrum itself should be regarded as one of the strongest proof that the Sun is condensed matter, as its generation requires a lattice which dictates the interatomic spacing of condensed matter.

It appears that the solar spectrum is reporting only a small fraction of the true energy content of the photosphere, providing information which is limited to the vibrational state of the solar lattice. Much more substantial energy is stored in the translational degrees of freedom. This is manifested by the convection currents of the granules [23] and the differential solar rotation observed by Carrington [18]. Moreover, there is strong evidence to suggest that sunspots are metallic [23] and, therefore, maintain electronic conduction bands with their own associated energy.

These realities explain why the temperature of the solar photosphere does not constitute a violation of the second law of thermodynamics. The 5,780 K [1, p.10] measured is linked only to the vibrational degrees of freedom of the photospheric lattice. However, the true energy of the photosphere is dominated by its translational degrees of freedom. This helps to account for the production of X-rays in association with solar flares rupturing the photospheric surface [34]. When this occurs, we are likely to be monitoring some measure of the translational energy associated with the photosphere, as matter moves horizontally across the surface and collides orthogonally with the flare's vertical displacement of material. In a sense, the flare is providing resistance to the horizontal flow of matter on the photosphere. As surface matter collides with the flare, its energy is revealed and X-ray emissions are obtained [34]. Similarly, the temperatures of the corona in the 2,000,000–3,000,000 K range [35, p.3–10] reflect a coupling of these atoms to the translational degrees of freedom on the photosphere. No violation of the second law exists. The energy content of the photosphere is likely to correspond to temperatures of ~7,000,000 K, when properly accounting for all of these phenomena as the author has previously stated [36]. In that case, the photospheric spectrum may be considered as reporting an apparent temperature, with little relevance to the real temperature of the surface [36]. Alternatively, it is also possible to reconcile the emission spectrum to the real temperature of the photosphere. The approach would be similar to that adopted when dealing with the microwave background problem [37] and, unfortunately, involves a reconsideration of Boltzmann's constant [38].

The consideration of energy partition in the Sun opens new avenues of discovery in physics. Most notably, it brings into question the universality of blackbody radiation, as first advocated by Gustav Kirchhoff [15,16]. *A priori*, the gaseous Sun fails to meet Kirchhoff's requirement for thermal equilibrium with an enclosure, as Max Planck recognized [22, §101]. Regrettably, Kirchhoff's law itself is unsound [39,40], destroying any perceived ability of gases to emit blackbody spectra. The issue is critical to the survival of the gaseous solar models. If local thermal equilibrium and its extension of Kirchhoff's formulation fails to guarantee that a blackbody spectrum is produced at the center of the Sun, then the gaseous models have no mechanism to generate its continuous emission. In part, this forms the basis of the solar opacity problem [31].

Dedication

This work is dedicated to the memory of Professor David G. Cornwell (10/8/1927–3/23/2011).

Submitted on July 2, 2011 / Accepted on July 9, 2011
First published online on July 16, 2011

References

1. Lang K.R. The Cambridge encyclopedia of the Sun. Cambridge University Press, Cambridge, 2001.
2. Clausius R. Über die bewegende Kraft der Wärme — Part I and Part II. *Poggendorfs Annalen der Physik und Chemie*, 1850, v.79, 368–397 and 500–524 (also in English: On the moving force of heat, and the laws regarding the nature of heat itself which are deducible therefrom. *Phil. Mag.*, 1851, v.2, 1–21 and 102–119).
3. Clausius R. The Mechanical Theory of Heat — with Its Applications to the Steam Engine and to Physical Properties of Bodies. John van Voorst, London. 1865.
4. Thomson W. On the dynamical theory of heat; with numerical results deduced from Mr. Joule's equivalent of a thermal unit and M. Regnault's observations on steam. *Math. and Phys. Papers*, 1851, v.1, 17583.
5. Wien W. Über die Energieverteilung in Emissionspektrum eines schwarzen Körpers. *Annalen der Physik*, 1896, v.58, 662–669.
6. Stefan J. Über die Beziehung zwischen der Warmestrahlung und der Temperature. *Sitzungsberichte der mathematischnaturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, Wien 1879, v.79, 391–428.
7. 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).
8. Langley S.P. Experimental determination of wave-lengths in the invisible spectrum. *Mem. Natl. Acad. Sci.*, 1883, v.2, 14762.
9. Langley S.P. On hitherto unrecognized wave-lengths. *Phil. Mag.*, 1886, v.22, 14973.
10. Langley S.P. The invisible solar and lunar spectrum. *Am. J. Science*, 1888, v.36(216), 39710.

11. Clerke A.M. A popular history of astronomy during the nineteenth century. Adam & Charles Black, London, 1893.
12. Meadows A.J. Early solar physics. Pergamon Press, Oxford, 1970.
13. Scheiner J. The temperature of the Sun. I. *Astron. Soc. Pac.*, 1898, v.9(64), 167–179.
14. Rogovsky E. On the temperature and composition of the atmosphere of the planets and the Sun. *Astrophys. J.*, 1901, v.14, 234–260.
15. 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.
16. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsvermögen und dem Absorptionsvermögen. der Körper für Wärme und Licht. *Poggendorff's 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).
17. Ebert H. On the electromagnetic nature of the solar radiation and on a new determination of the temperature of the Sun. *Astrophys. J.*, 1895, v.2, 55–57.
18. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v.3, 3–25 — A paper published in this Special Issue.
19. Robitaille P.M. A thermodynamic history of the solar constitution — II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v.3, 41–59 — A paper published in this Special Issue.
20. Very F.W. The temperature of the Sun. *Science*, 1908, v.27(685), 267–269.
21. Robitaille P.M. Max Karl Ernst Ludwig Planck (1858–1947). *Progr. Phys.*, 2007, v.4, 117–120.
22. Planck M. The Theory of Heat Radiation. P. Blakiston's Son & Co., Philadelphia, PA, 1914.
23. Robitaille P.M. On Solar Granulations, Limb Darkening, and Sunspots: Brief Insights in Remembrance of Father Angelo Secchi. *Progr. Phys.*, 2011, v.3, in press — A paper published in this Special Issue.
24. Robitaille P.M. The little heat engine: Heat transfer in solids, liquids, and gases. *Progr. Phys.*, 2007, v.4, 25–33.
25. Modest M.F. Radiative Heat Transfer. McGraw-Hill, New York, 1993, p.92–108.
26. Thirumaleshwar M. Fundamentals of Heat and Mass Transfer. Dorling Kindersley, Dehli, 2009, p.652.
27. Incropera F.P., DeWitt D.P., Bergman T.L., Lavine A.S. Fundamentals of Heat and Mass Transfer. 6th Edition, John Wiley & Sons, Hoboken, NJ, 2007.
28. Touloukian Y.S. and Ho C.Y. Thermophysical Properties of Matter (vols. 1). Plenum, New York, 1970.
29. Pierson H.O. Handbook of Carbon, Graphite, Diamond and Fullerenes. Noyes Publications, Park Ridge, NJ, 1993.
30. Haynes W.M. CRC Handbook of Chemistry and Physics. 91st Edition, CRC Press, Boca Raton, Fl, 2010–2011.
31. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v.3 — A paper published in this Special Issue.
32. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v.3, 60–74 — A paper published in this Special Issue.
33. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v.3, 764–770.
34. Kosovichev A. G., Zharkova V.V. X-ray flare sparks quake inside the Sun. *Nature*, 1998, v.393, 317–318.
35. Livingston W. and Koutchmy S. Eclipse Science Results: Past and Present. *ASP Conference Series*, v.205, 2000.
36. Robitaille P.-M. L. The collapse of the Big Bang and the gaseous Sun. *New York Times*, March 17, 2002 (available online: http://www.thermalphysics.org/pdf/times.pdf).
37. Penzias A. A. and Wilson R.W. A measurement of excess antenna temperature at 4080 Mc/s. *Astrophys. J.*, 1965, v.1, 419–421.
38. Robitaille P.M. Blackbody Radiation and the Loss of Universality: Implications for Planck's Formulation and Boltzman's Constant. *Progr. Phys.*, 2009, v.4, 14–16.
39. Robitaille P. M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v.3, 36–55.
40. Robitaille P. M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v.4, 3–13.

Stellar Opacity: The Achilles' Heel of the Gaseous Sun

Pierre-Marie Robitaille
Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA
E-mail: robitaille.1@osu.edu

The standard gaseous model of the Sun is grounded on the concept of local thermal equilibrium. Given this condition, Arthur Milne postulated that Kirchhoff's law could be applied within the deep solar interior and that a blackbody spectrum could be generated in this region, based solely on equilibrium arguments. Varying internal solar opacity then ensured that a blackbody spectrum could be emitted at the photosphere. In this work, it is demonstrated that local thermal equilibrium and solar opacity arguments provide a weak framework to account for the production of the thermal spectrum. The problems are numerous, including: 1) the validity of Kirchhoff's formulation, 2) the soundness of local thermal equilibrium arguments, 3) the requirements for understanding the elemental composition of the Sun, and 4) the computation of solar opacities. The OPAL calculations and the Opacity Project will be briefly introduced. These represent modern approaches to the thermal emission of stars. As a whole, this treatment emphasizes the dramatic steps undertaken to explain the origins of the continuous solar spectrum in the context of a gaseous Sun.

1 Introduction

The mechanism by which the solar spectrum is produced has long preoccupied astrophysics [1–4]. Though Langley established that the photosphere's emission [5–7] generally conformed to a blackbody lineshape [8,9], two lines of reasoning initially prevailed as to its formation. It was hypothesized that the photosphere contained condensed carbon [1, 2], as graphite was the premier blackbody source on Earth [3, 4]. Alternatively, it was believed that the pressure broadening of hydrogen could account for the spectrum [1, 2]. Although Kirchhoff had formulated his law of thermal emission in 1859 [10], observational astronomers appeared dissatisfied with the idea that Langley's spectrum [5–7] could be produced by assuming thermodynamic equilibrium and enclosure [9, p.1–45]. They insisted on placing carbon particles on the Sun for sixty years [1,2] and essentially dismissed any notion that Kirchhoff's law afforded a sufficient framework to generate the solar spectrum.

It would take the work of men [11] like Schuster [12], Schwarzschild [13], Eddington [14–17], Rosseland [18, 19] and Milne [20–23] to finally remove graphite from the Sun [2]. These communications [12–23] formed the foundation of radiation transfer within stars. They consequently came to represent the heart of modern stellar physics. As a group, these authors used elegant approaches, but without exception [12–23], their mathematical treatments relied on thermal equilibrium and the validity of Kirchhoff's law [10]. In addition, since the standard model of the Sun was deprived of condensed matter, astronomers would have to account for the production of the solar spectrum with physical atoms, ions, and electrons. Graphite was gone, but the theoretical alternative, solar opacity arguments, provided a questionable replacement.

2 Kirchhoff's law and local thermal equilibrium

Arthur Milne [2] was perhaps the first to advocate that the interior of the Sun could be regarded as existing in a state of local thermal equilibrium [20–23]. Milne's definition became central to astrophysical thought and will, therefore, be largely recalled: "*It is convenient to have a phrase to describe the circumstances under which the relation* $j_\nu = k_\nu B_\nu(T)$ *holds exactly. When a small portion of matter has a definite temperature* T*, and is behaving, i.e. emitting, as if it formed a part of an equilibrium enclosure at temperature* T*, we shall say that it is in "local thermodynamic equilibrium" at temperature* T*. We shall examine later in particular cases the conditions under which material is in local thermodynamic equilibrium. It is not necessary that the temperature shall be uniform. In an non-isothermal state, we may still have local thermodynamic equilibrium everywhere. The temperature may vary from point to point, but each point may be characterized by a definite temperature* T *and the element of matter at each point may be behaving as if in thermodynamic equilibrium at temperature* T" [23, p.81]. Milne's treatment was centered on Kirchhoff's law: $j_\nu = k_\nu B_\nu(T)$ [10]. Nonetheless, there was a risk that Milne's setting was so broad that virtually any non-equilibrium process, no matter how violent, could be considered in local thermal equilibrium, provided that sufficiently small volumes of matter were being considered. No restriction was placed on confirming the validity of these arguments.

Much like Milne, Chandrasekhar described local thermal equilibrium as follows: "*. . . we often encounter physical systems which, though they cannot be described as being in rigorous thermodynamical equilibrium, may yet permit the introduction of a temperature T to describe the local properties of the system to a very high degree of accuracy. The interior*

of a star, if in a steady and static state, is a case in point. For, even if the temperature at the center of the Sun, for instance, were 10^8 *degree, the mean temperature gradient would correspond to a change of only 6 degrees in the temperature over a distance of* 10^4 *cm. This fact, coupled with a probably high value for the stellar absorption coefficient, enables us to ascribe a temperature T at each point P such that the properties of an element of mass in the neighborhood of P are the same as if it were adiabatically inclosed in an inclosure at a temperature T*" [24, p.205]. Similar points were raised in Clayton's classic text [25, p.175]. These discussions were focused strongly on assumptions which pertain to a gaseous model.

On the surface, it would seem that Chandrasekhar's temperature gradient of only 6 degrees across 100 meters could be considered quite small [24, p.205]. Yet, the oceans of the Earth sustain convection currents based on much smaller temperature gradients. In fact, oceanographers might reject equilibrium arguments globally for the oceans, even though these temperature gradients are on the order of just a few degrees over spans of thousands of kilometers. The oceans contain convection currents as a direct manifestation of their lack of thermal equilibrium. Convection precludes the existence of equilibrium. As a result, a temperature variation of 6 degrees over a span of 100 meters should be treated as an enormous temperature gradient, not a condition approaching thermal equilibrium. The oceans demonstrate that Chandrasekhar's conditions, even if relaxed 1,000 fold, would still constitute powerful driving forces for convection, thereby eliminating all possibility of viewing the solar interior as existing in a state of thermal equilibrium.

Well before the days of Chandrasekhar, Milne elaborated further on local thermal equilibrium in the gaseous framework: "*The interior of a star is in a state of local thermodynamic equilibrium of this character. As we approach the boundary from the inside, the state of local thermodynamic equilibrium gives place to an entirely different state, in which the influence of external radiation on an element is paramount. It will be shown that when an element at temperature T is subjected to radiation, which is not black radiation of temperature T, the extent to which it behalves as if in thermodynamic equilibrium locally depends on the relative importance of collisions as a cause of atomic absorptions and emissions. If the atoms are sufficiently battered about by colliding with one another, they assume a state (distribution of stationary states) characteristic of thermodynamic equilibrium at temperature T; if they are not sufficiently battered about, their "temperature" becomes irrelevant and they emit and absorb at a rate which is determined by the incident radiation. It is clear that collisions will be the more numerous, and therefore likely to be more effective, the higher the density. This permits us to see in a general way why the state of local thermodynamic equilibrium in the interior of a star breaks down as we approach the surface...This assumption will certainly be satisfied in the far interior, since in the limit at great distances the conditions are those of an enclosure... It follows that the intensity of radiation at* $d\sigma$ *in the direction* θ *is* $B_\nu(T)$*, the intensity of black radiation for temperature T*" [23, p.81–83].

The argument advanced by Milne was framed in the context of the laws of gases. Milne saw the rapid collisions occurring at the center of the Sun as sufficient to establish equilibrium, but the requirements set forth by Kirchhoff [10] and Planck [8, 9] required something more significant. They demanded that the walls of the enclosure be rigid [9].

If a gas is highly compressed, the collisions with neighboring particles will enable the flow of heat through conduction. Gold has a density of 19.3 g/cm^3 [26, p.12–205] and many solids [26, p.12–80] have densities which are just slightly more than one order of magnitude (about a factor of 30) below the 150 g/cm^3 currently hypothesized for the center of the Sun [27, p.10]. When heat enters solids, it can travel through conduction, either thermally through its vibrational lattice or electronically through its conduction bands. Clearly, gases cannot sustain conduction bands, but they are subject to thermal conductive processes, especially at these densities. As such, when an atom in the gaseous model vibrates at the center of the Sun, it can transfer its energy to its "*non-rigid*" neighbor. Milne cannot assume that the atoms at the center of the Sun are devoid of collisional energy exchange, precisely because the atoms are not rigid. The center of the Sun cannot meet the requirements for a rigid enclosure as set forth by Kirchhoff and Planck [8–10]. The arguments of enclosure and "*local thermal equilibrium*" are invalid based on these considerations.

At the same time, Planck required that the source of blackbody radiation was found in material particles. Planck's entire *Heat Radiation* [9] was based on the analysis of a material oscillator not present at the center of the gaseous Sun: "*For among all conceivable distributions of energy the normal one, that is, the one peculiar to black radiation, is characterized by the fact that in it the rays of all frequencies have the same temperature. But the temperature of a radiation cannot be determined unless it be brought into thermodynamic equilibrium with a systems of molecules or oscillators, the temperature of which is known from other sources. For if we did not consider any emitting and absorbing matter there would be no possibility of defining the entropy and temperature of the radiation, and the simple propagation of free radiation would be a reversible process, in which the entropy and temperature of separate pencils would not undergo any change. Now we have deduced in the preceding section all the characteristic properties of the thermodynamic equilibrium of a system of ideal oscillators. Hence, if we succeed in indicating a state of radiation which is in thermodynamic equilibrium with the system of oscillators, the temperature of the radiation can be no other than that of the oscillators, and therewith the problem is solved*" [9, §144].

Max Planck required that a perfect absorber be present in order to produce blackbody radiation. Milne neglected this important line from *Heat Radiation*: "*Hence in a vacuum bounded by totally reflecting walls any state of radiation may persist*" [9, §51]. Planck then argued that, if an arbitrarily small quantity of matter was introduced, the radiation in the enclosure will change to a new state. However, it will not be a blackbody state unless the substance is not transparent for any frequency. Planck chose a piece of carbon to ensure blackbody radiation [9, §51]. The desired radiation does not simply appear [9, §51], as Milne and his contemporaries surmised. The presence of an enclosure, by itself, could never satisfy the requirements for the production of blackbody radiation. Planck insisted throughout *Heat Radiation* on the need for a physical oscillator and he reminded his readers that only "*material particles*" can be involved in emission [9, §4] and absorption [9, §12]. A physical oscillator which acted as a perfect absorber must be present. Milne has not advanced such a species at the center of the Sun.

Instead, Milne, like Schuster [12], Schwarzschild [13], and Eddington [14–17] before him, automatically presumed that the invocation of Kirchhoff's law provided sufficient proof that the interior of the Sun harbored black radiation, despite the absence of the rigid enclosure required by Kirchhoff [10]. Blackbody radiation was inserted at the center of the Sun without any requirement on the material generating the needed photons. All that was required was enclosure (even if not strictly rigid) and a newly hypothesized "*local thermodynamic equilibrium*". For Milne, the presence of an enclosure was insured by the hypothesis that the density at the center of the Sun was sufficiently elevated to restrict photonic and atomic diffusion [20–23].

In reality, Milne's idea fell far short of the requirements to produce blackbody radiation. He was considering a setting where conduction, not radiation, could dominate heat exchange. Consequently, his arguments relative to radiative heat transfer were without strong scientific justification. Milne had neglected the observation that the collision of adjacent atoms constituted the universally accepted exchange mechanism for thermal conduction, not equilibrium. It was for this reason that Planck insisted on a rigid enclosure.

A careful review of blackbody radiation has revealed that the production of such a spectrum always requires the presence of a perfect absorber [3]. Planck himself constantly brought forth the carbon particle as inherently linked to the validity of his arguments [3]. Kirchhoff's reasoning that an adiabatic enclosure could contain black radiation has been exposed as flawed and his law of thermal emission as erroneous [3, 4, 28–30]. The universality of blackbody radiation simply does not exist [3, 4, 28–30]. Yet, even if Kirchhoff's law was valid, Milne's argument was fallacious, as he lacked both the rigid enclosure and the materially perfect oscillator required by Max Planck to ensure that a blackbody spectrum could be produced at the center of the Sun.

3 Solar and stellar opacity

Solar opacity [22, 31, 32, 34–39] plays a vital role in all modern gaseous models of the Sun [24, 25, 40–46] and is currently at the center of our understanding of the stars. Therefore, the study of solar opacity has far reaching implications throughout modern astronomy.

Opacity, κ, refers to the ability of a material to absorb incoming radiation. Monochromatic opacity, κ_ν, is associated with a single frequency. The extinction coefficient, α (cm^{-1}), is equal to the opacity, κ (cm^2/g) multiplied by the density of the material, ρ (g/cm^3).

To calculate opacity within the solar interior, solar physicists first accept that the Sun can radiate internally. By itself, this constitutes a notable departure from the rest of Earthly physics. For all objects on Earth, internal heat transfer occurs through conductive and convective paths, not internal radiation. Radiation allows objects to achieve thermal equilibrium with one another, not within themselves. As a result, the idea that the Sun transfers internal energy through radiation directly implies that astrophysics treats the solar interior as the sum of its individual atomic, ionic, and electronic species. The Sun as a single object does not exist in the gaseous models. Only in such a scenario would internal radiation permit the transfer of energy between the constituent objects which make up the Sun. Still, Milne required that, within the center of the Sun, atoms, ions, and electrons were packed such that collisions occur. This scenario rendered conduction probable, greatly impacting any radiative field.

In gaseous solar models, thermal photons at X-ray frequencies, with a characteristic blackbody appearance, are believed to be produced at the center of the Sun. Over the course of thousands of years, Eddington stated that these thermal photons slowly leaked out of the solar body [16]. As they traversed increasingly elevated layers of the solar mass, photons gradually lost some of their energy. The entire solar spectrum was shifting from the X-ray to the visible range, while preserving a blackbody appearance [16].

3.1 Opacity mechanisms

Stellar opacity involves the removal of energy from a beam of photons originating in the core of the Sun through four mechanisms: 1) bound-bound, 2) bound-free, 3) free-free, and 4) scattering processes (see [41, p.137–141] for an excellent description). Bound-bound processes rely on spectroscopic line absorption, either within an atom or an ion. Bound-free mechanisms result in the dissociation of a previously bound electron by an incoming photon. The electron becomes completely free of the atom or ion. Free-free processes are inverse Bremsstrahlung mechanisms, whereby a free electron and an ion interact during which time the combined species is able to absorb a photon [41, p.138]. In scattering mechanisms, the momentum of the photon is being transferred to a scattering electron. Theoretical astrophysics

calculates opacities for the Sun by taking the summation of these processes, for all atoms, ions, and electrons at all temperatures within the solar interior.

The negative hydrogen ion was advanced as a significant determinant of solar opacity by Wildt [47]. The concept immediately received the support of Chandrasekhar who calculated that the negative hydrogen atom within the context of a gaseous solar model would contribute greatly to solar opacity in the 4,000–24,000 Å range [48–51]. Of course, the negative hydrogen ion spectrum extended over much of the photospheric emission (~2,500–25,000 Å).

Nonetheless, the negative hydrogen ion could never, by itself, generate the continuous solar spectrum with its characteristic thermal appearance. For gaseous models, the production of the thermal spectrum involves the slow conversion of a hypothetically X-ray blackbody spectrum produced in the solar interior to the visible spectrum observed at the photosphere. Thus, if a blackbody spectrum did exist at the center of the Sun, it would be characterized by a Wien displacement temperature of ~15,000,000 K. Such a spectrum would be centered in the X-ray region. It would then have to be gradually shifted, while always maintaining its thermal appearance, to much lower frequencies.

Consequently, astrophysics is requiring that a perfect mixture of atoms, ions, and electrons exists at all layers within the Sun. In each layer, these mixtures could then produce the desired local blackbody spectrum. Within each solar layer, a new perfect mixture must exist in order that its absorptive characteristics enable the production of a new shifted thermal spectrum.

Therefore, despite Chandrasekhar's findings [48–51], the computation of solar opacity has remained a tremendously complex undertaking. For example, the American astrophysics community has invested heavily in calculating the opacity contributions from neutral and ionized gases. In a project involving international collaboration, the Los Alamos National Laboratory led Opacity Project [33, 34] provided an absolutely phenomenal treatment of nearly every possible atomic species inside the stars, in widely varying states of oxidation. Similar findings have been obtained at the Lawrence Livermore National Laboratories. These studies have resulted in the OPAL opacity values [35–39], but none of the opacity mechanisms considered by these methods can be used to explain the origin of the blackbody spectrum in graphite. This suggests that these mechanisms are not truly related to the production of the solar spectrum.

3.2 Rosseland mean opacities

The determination of internal solar opacity values must be performed at each individual frequency of interest, since the production of a blackbody spectrum always remains frequency dependent. The problem becomes so overwhelming that astrophysics has chosen to adopt Rosseland mean opacities [18, 19]. Through Rosseland's approach, a single frequency independent value of opacity can be obtained for each solar level.

On the surface, it could be argued that Rosseland mean opacities merely reduce an otherwise intractable problem. They lower computational requirements and greatly simplify the presentation of opacity data. Rosseland mean opacities enable solar physics to sidestep the reality that, at each level of the solar interior, it is impossible to generate a purely blackbody spectrum with strict adherence to Planckian behavior at all frequencies. It is not feasible to build a blackbody spectrum from the sum of non-blackbody processes. For instance, during the computation stage, a single bound-bound transition will introduce a "*spike-like*" contribution in the calculated spectrum. Each "*spike*" being associated with line absorption. Such a "*spike*" must then be compensated by using the sums of processes (other bound-bound processes, or bound-free, free-free and scattering mechanisms) whose existence will always remain in doubt at the levels required to incorporate the initial "*spike*" into the final solution for the blackbody lineshape. The entire process becomes an exercise in parameter fitting, devoid of confirmatory physical evidence.

Still, Rosseland mean opacities remain at the heart of modern solar models [24, 25, 40–46]. Within each layer in the Sun, a mean opacity can be inferred based on expected atomic, ionic, and electronic species. However, the sum of the processes (bound-bound, bound-free, free-free, scattering) utilized in Rosseland mean opacity computations cannot be infinite. Thus, rather than analyze mean opacities, scientists can convince themselves of the futility of these approaches by taking the mean opacity solutions and using the same species and concentrations to calculate the associated *frequency dependent spectra*. Such solutions will not correspond to black body spectra. As a result, Rosseland mean opacities form a weak foundation for the gaseous solar models. The summation of numerous spectral processes which are individually unrelated to thermal radiation can never give rise to a truly black spectrum.

3.3 Elemental compositions

To further complicate matters, the computation of solar opacity, as a function of depth, requires that the elemental composition of the Sun [52] remains independent of spatial position. Such, a requirement can never be justified. Our current understanding of the solar composition rests, and will always rest, on that which can be evaluated at the level of the photosphere. All extensions of the solar composition to the solar interior and all claims of constant elemental constitution with depth should be regarded as scientific conjecture.

4 Conclusion

Through opacity considerations, solar physicists believe that an X-ray based blackbody spectrum, produced at the center of the Sun, can be emitted at the solar surface in the visible

range. However, from the moment that the Sun was hypothesized to exist in the gaseous state in the mid-1800s, objections were raised as to the ability of gases to emit a blackbody spectrum [1]. The interior of a gaseous Sun was thought to be essentially transparent to radiation. This was the position advocated by Herbert Spencer when he complained that, if sunspots were openings in the photosphere, one should be able to see through them to the other side [1]. In fact, the same "*famous objection*" was voiced by Kirchhoff himself [1]. According to Kirchhoff, the interior of the Sun could only sustain blackbody radiation if it was surrounded by a condensed photosphere [1]. Kirchhoff well understood that no gas, in isolation, ever produced a blackbody spectrum. The presence of condensed matter was always required.

In support of Kirchhoff's liquid photosphere [1], there are numerous lines of evidence that the photosphere is condensed matter [53]. Granules, sunspots, and limb darkening provide additional evidence [56]. Sunspot emissivities are highly suggestive of metallic character [56] strengthening the case for condensed matter. All of these factors should be considered when advancing the proper phase of the photosphere and the mechanism associated with solar thermal emission.

Nonetheless, despite clear violations with regards to enclosure, thermal equilibrium, and the presence of a perfect absorber as required by Max Planck [9], solar physics has tried to account for the generation of the Planckian spectrum. Yet, none of the mechanisms advanced can be used to explained the simple thermal spectrum of graphite itself. In fact, although physics advocates an understanding of internal thermal radiation within the Sun, it has produced no mechanism by which the simplest earthly spectrum can be explained. This constitutes a powerful reminder that tremendous difficulties remain relative to the science of blackbody radiation [3, 4]. In the end, stellar opacity calculations represent a myriad of physical impossibilities. None of the suggested opacity mechanisms (bound-bound, bound-free, free-free, and scattering) are related to the emission of a single photon by graphite.

As such, beyond an inability to support structure, the shortcomings of any gaseous solar model rests on opacity. Even though Milne and his predecessors were incorrect in inferring that a blackbody spectrum could be produced at the center of the Sun, the gaseous models contain numerous other stumbling blocks on their way to generating a continuous spectrum at the solar surface. A truly remarkable thesis has been advanced to explain the photospheric spectrum within the gaseous model. In the end, astrophysics has championed a solution for obtaining the solar spectrum which cannot survive the careful scrutiny of the spectroscopic scientific community.

Each spectroscopic signature in nature is linked to a unique physical process. For instance, a Lyman or a Balmer series can only be produced by electronic transitions within the hydrogen atom. Similarly, atomic line spectra are unique to each individual elemental or ionic species. Nuclear magnetic resonance (NMR) spectra are obtained from particular spin transitions within a well defined physical and experimental context. Physics does not search for the Lyman series in NMR spectra. One process is electronic, the other nuclear. Within the gaseous Sun, modern astrophysics currently believes that it can produce the graphitic spectrum using processes which do not exist in graphite. It is improper to advance that a blackbody spectrum can be produced in the Sun using physical mechanism which are not present on Earth within all the blackbodies currently studied in our laboratories [3, 4]. The use of a nearly infinite sum of atomic, ionic, and electronic processes which can alter their absorption and emission precisely in a manner which preserves the blackbody appearance of the solar spectrum at all depths within the Sun represents a non-scientific exercise based solely on the desire to salvage the gaseous equations of state. It is well-known that thermal emissivity in gases can drop with increasing temperature. Neither pressure broadened gases nor any of the atomic, ionic, and electronic processes advocated in the interior of the Sun have a fourth power of temperature behavior. Furthermore, the gaseous models depend on knowledge of the internal constitution of the stars based on the solar elemental constituents. Mankind will always lack such information.

As a result, this work constitutes an invitation to reconsider the phase of the Sun [53–55]. The gaseous models suffer from two insurmountable weaknesses: 1) the inability to account for photospheric structures [56], and 2) the lack of a proper mechanism to generate the solar spectrum. Observational astrophysics has long documented the existence of features of the solar surface which demand the presence of condensed matter [56]. The belief that opacity arguments can account for the illusionary nature of the solar surface and all associated structures, discounts the realization that the photosphere also *behaves* as condensed matter [56, 57]. Helioseismology demonstrates that the Sun acts as a resonant cavity [53]. On Earth, resonant cavities are manufactured from condensed matter [4]. It is not reasonable to expect that a gaseous Sun can create an illusionary surface in the visible range using negative hydrogen ion opacity, while at the same time and in the same layer, produce a surface which is nearly perfectly reflecting for wavelengths which extend over many thousands of meters. Such are the requirements, if the Sun really acts as a resonant cavity [58, p.60]. Perfect resonators sustain standing waves which are never absorbed [4]. Accordingly, the photosphere of the gaseous Sun must be strongly opaque in the visible region while powerfully reflecting in the sub-audio. In addition, the gaseous models must account for the presence of transverse waves on the surface of the Sun when gases are known to sustain only longitudinal waves [53, 57]. It remains the case that seismology is a science of condensed matter [53]. To account for seismological behavior in a gaseous Sun using opacity arguments consti-

tutes a significant departure from accepted Earthly physics.

Given the problems which surround solar opacity, it remains difficult to understand how the gaseous models of the Sun have survived over much of the twentieth century. Local thermal equilibrium does not exist at the center of the Sun. Both Kirchhoff and Planck require rigid enclosure which is not found in the Sun [9, 10]. Planck has also warned that the Sun fails to meet the requirements for being treated as a blackbody [59]. Milne's rapid collisional regime constitutes a path to conduction, not equilibrium [20–23]. Milne and his contemporaries cannot infer that a blackbody spectrum exists at the center of the Sun based on Kirchhhoff's law [10], even if the law was valid [60]. Unfortunately, not only does the Sun fail to meet the requirements for enclosure and local thermal equilibrium, but Kirchhoff's law itself is erroneous [3,4]. The production of a blackbody spectrum requires the presence of a perfect absorber. Max Planck appeared well-aware of this reality [3,59]. Gaseous opacity arguments will always fall far short of what was required. In the end, the mechanism used to generate the solar spectrum should be shared with graphite itself. The most likely physical cause remains the vibration of atomic nuclei within the confines of a layered graphite-like lattice [28, 55].

Dedication

This work is dedicated to my sister, Lydia, and to her children, Fabienne and Louis.

Submitted on July 16, 2011 / Accepted on July 19, 2011
First published online on July 26, 2011

References

1. Robitaille P.-M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v.3, 3–25 — A paper published in this Special Issue.
2. Robitaille P.-M. A thermodynamic history of the solar constitution — II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v.3, 41–59 — A paper published in this Special Issue.
3. Robitaille P.-M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v.3, 36–55.
4. Robitaille P.-M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v.4, 3–13.
5. Langley S.P. Experimental determination of wave-lengths in the invisible spectrum. *Mem. Natl. Acad. Sci.*, 1883, v.2, 147–162.
6. Langley S.P. On hitherto unrecognized wave-lengths. *Phil. Mag.*, 1886, v.22, 149–173.
7. Langley S.P. The invisible solar and lunar spectrum. *Am. J. Science*, 1888, v.36 (216), 397–410.
8. Planck M. Ueber das Gesetz der Energieverteilung in Normalspectrum. *Annalen der Physik*, 1901, v.4, 553–563.
9. Planck M. The Theory of Heat Radiation. P. Blakiston's Son & Co., Philadelphia, PA, 1914.
10. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsverögen und dem Absorptionsvermögen. der Körper für 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.*, ser. 4, 1860, v.20, 1–21).
11. Menzel D.H. Selected Papers on the Transfer of Radiation: 6 papers by Arthur Schuster, K. Schwarzschild, A.S. Eddington, S. Rosseland, E.A. Milne. Dover Publications, Inc., New York, 1966.
12. Schuster A. Radiation through a foggy atmosphere. *Astrophys. J.*, 1905, v.21, 1–22 (also found in [11], p.3–24).
13. Schwarzschild K. Über das Gleichgewicht der Sonnenamospäre. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttigen, Göttinger Nachrichten*, 1906, 195, 41–53 (English translation: Concerning the equilibrium of the solar atmosphere, found in [11], p.25–34 and Meadows A.J. Early Solar Physics. Pergamon Press, Oxford, 1970., p.277–290).
14. Eddington A.S. On the radiative equilibrium of the stars. *Mon. Not. Roy. Astron. Soc.*, 1916, v.77(1), 16–35 (Also found in Lang K.R. and Gingerich O.: A Source Book in Astronomy and Astrophysics, 1900-1975. Harvard University Press, Cambridge, MA, 1979, p.225–235; also found in [11], p.53–72).
15. Eddington A.S. The internal constitution of the stars. *The Observatory*, 1920, v.43 (557), 341–357.
16. Eddington A.S. The Internal Constitution of the Stars. Dover Publ. Inc., New York, 1959.
17. Eddington A.S. Stars and Atoms. Yale University Press, New Haven, 1927.
18. Rosseland S. Note on the absorption of radiation within a star. *Mon. Not. Roy. Astron. Soc.*, v.84(7), 525–528 (also found in [11], p.73–76).
19. Rosseland S. Theoretical Astrophysics: Atomic Theory and the Analysis of Stellar Atmospheres and Envelopes. Clarendon Press, Oxford, 1939.
20. Milne E.A. Selective radiation-pressure and the structure of a stellar atmosphere. *Mon. Not. Roy. Astron. Soc.*, 1927, v.87, 697–708.
21. Milne E.A. The effect of collisions on monochromatic radiative equilibrium. *Mon. Not. Roy. Astron. Soc.*, 1928, v.88, 493–502.
22. Milne E.A. Bakerian Lecture: The structure and opacity of a stellar atmosphere. *Phil. Trans. Roy. Soc. London*, 1929, v.228, 421–461.
23. Milne E.A. Thermodynamics of the stars. *Handbuch der Astrophysik*, 1930, v.3, Part 1, 65–255 (also in [11], p. 77–269).
24. Chandrasekhar S. Introduction to the Study of Stellar Structure (Reprint of the 1939 Edition). Dover Publications, Inc., 1957.
25. Clayton D.D. Principles of Stellar Evolution and Nucleosynthesis. McGraw-Hill, New York, 1968.
26. Haynes W.M. CRC Handbook of Chemistry and Physics. 91st Edition, CRC Press, Boca Raton, Fl, 2010–2011.
27. Lang K.R. The Cambridge Encyclopedia of the Sun. Cambridge University Press, Cambridge, 2001.
28. Robitaille P.-M. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v.31(6), 1263–1267.
29. Robitaille P.-M. An analysis of universality in blackbody radiation. *Progr. Phys.*, 2006, v.2, 22–23; arXiv: physics/0507007.
30. Robitaille P.-M. A critical analysis of universality and Kirchhoff's law: a return to Stewart's law of thermal emission. *Progr. Phys.*, 2008, v.3, 30–35; arXiv: 0805.1625v1 [physics.gen-ph].
31. Carson T.R. Stellar opacity. *Ann. Rev. Astron. Astrophys.*, 1976, v.14, 95–117.
32. Rozsnyai B.F. Solar opacities. *J. Quant. Spec. Rad. Trans.*, 2001, v.71, 655–663.
33. The Opacity Project Team. The Opacity Project. Institute of Physics Publishing, Bristol, UK, 1995, v.1.
34. The Opacity Project Team. The Opacity Project. Institute of Physics Publishing, Bristol, UK, 1996, v.2.

35. Rogers F.J. and Iglesias C.A. Stellar opacity. Lawrence Livermore National Laboratory, Preprint UCLRL-JC-137066.
36. Iglesias C.A. and Rogers F.J. Opacities for the solar radiative interior. *Astrophys. J.*, 1991, v.371, 408–417.
37. Iglesias C.A. and Rogers F.J. Radiative opacities for carbon- and oxygen-rich mixtures. *Astrophys. J.*, 1993, v.412, 752–760.
38. Rogers F.J. and Iglesias C.A. Rosseland mean opacities for variable compositions. *Astrophys. J.*, 1992, v.401, 361–366.
39. Iglesias C.A. and Rogers F.J. Updated OPAL opacities. *Astrophys. J.*, 1996, v.464, 943–953.
40. Reddish V.C. The Physics of Stellar Interiors. Edinburgh University Press, Edinburgh, 1974.
41. Kippenhahn R. and Weigert A. Stellar Structure and Evolution. Springer-Verlag, Berlin, 1994.
42. Bhatnagar A. and Livingston W. Fundamentals of Solar Astronomy. World Scientific Series in Astronomy and Astrophysics — v.6, World Scientific, New Jersey, 2005.
43. Bahcall J.N. and Pinsonneault M.H. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.*, 1992, v.64, 885–926.
44. Bahcall J.N. and Pinsonneault M.H. What do we (not) know theoretically about solar neutrino fluxes? *Phys. Rev. Letters*, 2004, v.92, 121301(1–4).
45. Bahcall J.N., Serenelli A.M. and Basu S. New solar opacities, abundances, helioseismology, and neutrino fluxes. *Astrophys. J.*, 2005, v.621, L85–L88.
46. Bahcall J.N., Serenelli A.M. and Pinsonneault M. How accurately can we calculate the depth of the solar convective zone? *Astrophys. J.*, 2004, v.614, 464–471.
47. Wildt R. Electron affinity in astrophysics. *Astrophys. J.*, 1939, v.89, 295–301.
48. Chandrasekhar S. On the continuous absorption coefficient of the negative hydrogen ion. *Astrophys. J.*, 1945, v.102, 223–231.
49. Chandrasekhar S. On the continuous absorption coefficient of the negative hydrogen ion. II. *Astrophys. J.*, 1945, v.102, 395–401.
50. Chandrasekhar S. and Breen F.H., On the continuous absorption coefficient of the negative hydrogen ion. III. *Astrophys. J.*, 1946, v.104, 430–445.
51. Chandrasekhar S. and Münch G. The continuous spectrum of the Sun and the stars. *Astrophys. J.*, 1946, v.104, 446–457.
52. Grevesse N. and Noels A. Cosmic abundances of the elements. In: *Origin and evolution of the elements* (N. Prantzos, E. Vangioni-Flam and M. Cassé, eds), Cambridge University Press, Cambridge, 1993.
53. Robitaille P.-M. The solar photosphere: evidence for condensed matter. *Progr. Phys.*, 2006, v.2, 17–21 (also found in slightly modified form within *Research Disclosure*, 2006, v.501, 31–34; title no.501019).
54. Robitaille P.-M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v.1, 70–81 (also in arXiv: astro-ph/0410075).
55. Robitaille P.-M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v.3, 60–74 — A paper published in this Special Issue.
56. Robitaille P.-M. On solar granulations, limb darkening, and sunspots: Brief insights in remembrance of Father Angelo Secchi. *Progr. Phys.*, 2011, v.3, 79–88 — A paper published in this Special Issue.
57. Robitaille P.-M. On the presence of a distinct solar surface: A reply to Hervé Faye. *Progr. Phys.*, 2011, v.3, 75–78 — A paper published in this Special Issue.
58. Muller R. The solar granulation. In: *Motions in the solar atmosphere* (A. Hanslmeier and M. Messerotti, eds.), Astrophysics and Space Science Library, v.239, Kluwer Academic Publishers, Dordrecht, 1999, p.35–70.
59. Robitaille P.-M. On the temperature of the photosphere: Energy partition in the Sun. *Progr. Phys.*, 2011, v.3, 89–92 — A paper published in this Special Issue.
60. Robitaille P.-M. The collapse of the Big Bang and the gaseous Sun. *New York Times*, March 17, 2002, page A10 (available online: http://thermalphysics.org/pdf/times.pdf).

Lessons from the Sun

Pierre-Marie Robitaille
Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA
E-mail: robitaille.1@osu.edu

In this brief note, the implications of a condensed Sun will be examined. A celestial body composed of liquid metallic hydrogen brings great promise to astronomy, relative to understanding thermal emission and solar structure. At the same time, as an incompressible liquid, a condensed Sun calls into question virtually everything which is currently believed with respect to the evolution and nature of the stars. Should the Sun be condensed, then neutron stars and white dwarfs will fail to reach the enormous densities they are currently believed to possess. Much of cosmology also falls into question, as the incompressibility of matter curtails any thought that a primordial atom once existed. Aging stars can no longer collapse and black holes will know no formative mechanism. A condensed Sun also hints that great strides must still be made in understanding the nature of liquids. The Sun has revealed that liquids possess a much greater potential for lattice order than previously believed. In addition, lessons may be gained with regards to the synthesis of liquid metallic hydrogen and the use of condensed matter as the basis for initiating fusion on Earth.

"*Young people, especially young women, often ask me for advice. Here it is, valeat quantum. Do not undertake a scientific career in quest of fame or money. There are easier and better ways to reach them. Undertake it only if nothing else will satisfy you; for nothing else is probably what you will receive. Your reward will be the widening of the horizon as you climb. And if you achieve that reward you will ask no other.*"

Cecilia Payne-Gaposchkin [1]

When Cecilia Payne [1] discovered that the stars are primarily composed of hydrogen [2], she encountered strong opposition from Arthur Eddington, her first mentor, and from Henry Norris Russell [3]. Nonetheless, Cecilia Payne's work engendered a new age in astronomy: hydrogen became the building block of the universe. Russell would eventually come to echo Payne's position [4]. In those days, it was natural to assume that a hydrogen-based Sun would be gaseous [5, 6]. Ten years after Payne published her classic report, Wigner and Huntington proposed that condensed metallic hydrogen could be synthesized [7]. In so doing, they unknowingly provided James Jeans with the material he had lacked in constructing liquid stars [5]. Still, though liquid metallic hydrogen became a component of the giant planets and the white dwarf [8], the concept of condensed matter was kept well removed from the Sun.

Now that liquid metallic hydrogen has been advanced as a solar building block (see [8] and citations therein), it is likely that opposition will be raised, for many will foresee unsettling changes in astronomy. A liquid Sun brings into question our understanding of nearly every facet of this science: from stellar structure and evolution [9], the existence of black holes [10], the premordial atom [11], dark energy [12], and dark matter [13]. It is not easy to abandon familiar ideas and begin anew.

However, some scientists will realize that a liquid metallic hydrogen model of the Sun [8], not only opens new avenues, but it also unifies much of human knowledge into a cohesive and elegant framework. A liquid metallic Sun invites astronomy to revisit the days of Kirchhoff [14] and Stewart [15], and to recall the powerful lessons learned from studying the thermal emission of materials [16,17]. It emphasizes that our telescopes observe structural realities and not illusions [18, 19]. In recognizing the full character of these structures, all of the great solar astronomers from Galileo [20], to Secchi [21], to Hale [22] are honored. These observers knew that solar structures (granules, sunspots, pores, flares, prominences, etc. . .) were manifesting something profound about nature.

For astrophysicists, the Sun imparts lessons which may well have direct applications for mankind. For instance, the solar body holds the key to fusion. If the Sun is made from condensed matter [8], then our experiments should focus on this state. Sunspots may also guard the secret to synthesizing metallic hydrogen on Earth [8]. If sunspots are truly metallic [18], as reflected by their magnetic fields [22], then attempts to form liquid metallic hydrogen on Earth [8] might benefit from the presence of magnetic fields. Our analysis of the photospheric constitution and the continuous thermal spectrum should be trying to tell us something about liquids and their long range order. It is currently believed that liquids possess only short term order [23]. In this regard, perhaps physics has lacked caution in bombarding the fragile liquid lattice with X-rays and neutrons [24, 25]. These methods may fail to properly sample the underlying structure. Gentler

approaches may reveal structure where none was previously believed to exist. The solar spectrum implies long range order, much like that observed in graphite [16, 17, 26]. As such, liquid metallic hydrogen on the photosphere could provide the framework for long range order, despite the fact that its only binding force lies in the need to maintain electronic conduction bands (see [8] and references therein). Most importantly, however, the Sun might be trying to tell us that we still do not properly understand thermal emission [16, 17, 27]. If gaseous models exist to this day, it is because the mechanism which produces the blackbody spectrum in graphite continues to be elusive [16, 17, 27]. Of all spectroscopic signatures, blackbody radiation remains the only one which has not been explained fully. These problems constitute serious and important questions for humanity. Unlocking these mysteries is certain to keep scientists occupied, as we continue to ponder upon the lessons discerned from the Sun.

Dedication

This work is dedicated to the memory of Miss Beckly [28, p.134], Annie Scott Dill Russell [28, p.144–146], Margaret Huggins [29], Henrietta Swan Leavitt [30, 31], Annie Jump Cannon [32–34], Antonia Maury [35], Williamina Paton Stevens Fleming [36–38], Cecilia Payne-Gaposchkin [1] and the forgotten women of astronomy [39, 40].

Submitted on July 23, 2011 / Accepted on July 24, 2011
First published in online on July 26, 2011

References

1. Gingerich O. Obituary — Payne-Gaposchkin Cecilia. *Quart. J. Roy. Astron. Soc.*, 1982, v.23, 450–451.
2. Payne C.H. The relative abundances of the elements. *Stellar Atmospheres*. Harvard Observatory Monograph no. 1, Harvard University Press, Cambridge, MA, 1925, Chapter 13 (reprinted in Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p.245–248).
3. Greenstein G. The ladies of Observatory Hill: Annie Jump Cannon and Cecilia Payne-Gaposchkin. *Am. Scholar*, 1993, v.62(3), 437–446.
4. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v.70, 11–82.
5. Robitaille P.-M. A Thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v.3, 3–25 — A paper published in this Special Issue.
6. Robitaille P.-M. A Thermodynamic history of the solar constitution — II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v.3, 41–59 — A paper published in this Special Issue.
7. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v.3, 764–770.
8. Robitaille P.-M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v.3, 60–74 — A paper published in this Special Issue.
9. Kippenhahn R. and Weigert A. Stellar Structure and Evolution. Springer-Verlag, Berlin, 1994.
10. Webster B.L. and Murdin P. Cygnus X-1: A spectroscopic binary with a heavy companion? *Nature*, 1972, v.235, 37–38 (reprinted in: Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p.462–463)
11. Lemaître G. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Ann. Société Scientifique de Bruxelles*, 1927, v.47, 49–59.
12. Frieman J.A., Turner M.S. and Huterer D. Dark energy and the accelerating universe. *Ann. Rev. Astron. Astrophys.*, 2008, v.46, 385–432.
13. Bertone G., Hooper D., Silk J. Particle Dark Matter: Evidence, candidates and constraints. *Phys. Rep.*, 2005, v.405, 279–390; arXiv: hep-ph/0404175.
14. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsverögen und dem Absorptionsvermögen. der Körper für Wärme und Licht. *Poggendorff's 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.*, ser. 4, 1860, v.20, 1–21).
15. 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(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).
16. Robitaille P.-M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v.3, 36–55.
17. Robitaille P.-M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v.4, 3–13.
18. Robitaille P.-M. On Solar granulations, limb darkening, and sunspots: Brief insights in remembrance of Father Angelo Secchi . *Progr. Phys.*, 2011, v.3, 79–88 — A paper published in this Special Issue.
19. Robitaille P.-M. On the presence of a distinct Solar surface: A reply to Hervé Faye. *Progr. Phys.*, 2011, v.3, 75–78 — A paper published in this Special Issue.
20. Galilei G. and Scheiner C. On Sunspots. Translated by E. Reeves and A.V. Helden, University of Chicago Press, Chicago, 2010.
21. Secchi A. Le Soleil. Gauthier-Villars, Paris, 1870.
22. Hale G.E. On the probable existence of a magnetic field in Sun-spots. *Astrophys. J.*, 1908, v.28, 315–343.
23. March N.H. and Tosi M.P. Introduction to Liquid State Physics. World Scientific, River Edge, NJ, 2002.
24. Fischer H.E., Barnes A.C. and Salmon P.S. Neutron and x-ray diffraction studies of liquids and glasses. *Rep. Prog. Phys.*, 2006, v.69, 233–299.
25. Leclercq-Hugeux F., Coulet M.-V., Gaspard J.-P., Pouget S. and Zanotti J.M. Neutrons probing the structure and dynamics of liquids. *Comptes Rendus Physique*, 2007, v.8(7–8), 884–908.
26. Pierson H.O. Handbook of Carbon, Graphite, Diamond and Fullerenes. Noyes Publications, Park Ridge, NJ, 1993.
27. Robitaille P.M.L. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v.31(6), 1263–1267.
28. Clark S. The Sun Kings: The Unexpected Tragedy of Richard Carrington and the Tale of How Modern Astronomy Began. Princeton University Press, Princeton, 2007.
29. Becker B.J. Dispelling the myth of the able assistant: Margaret and William Huggins at work in the Tulse Hill Observatory in *Creative Couples in the Sciences* (H.M. Pycior, N.G. Slack, and P.G. Abir-Am, eds.), 1996, Rutgers University Press, 88–111.
30. Pickering E.C. Periods of 25 Variable stars in the Small Magellanic Cloud. *Harvard College Observatory Circular*, 1912, v.173, 1–3.

31. Mitchell H.B. Henrietta Leavitt and the Cepheid variables. *The Physics Teacher*, 1976, v.14(3), 162–167.

32. Bok P.F. Annie Jump Cannon 1863–1941. *Publ. Astron. Soc. Pac.*, 1941, v.53, 168–170.

33. Merrill P. Annie Jump Cannon. *Mon. Not. Roy. Astron. Soc.*, 1942, v.102(2), 74–76.

34. Gaposchkin C.P. Annie Jump Cannon. *Science*, 1941, v.93(2419), 443–444.

35. Hoffleit D. Antonia C. Maury. *Sky & Telescope*, 1952, v.11(5), 106.

36. Cannon A.J. Williamina Paton Fleming, *Science*, 1911, v.33(861), 987–988.

37. Cannon A.J. Williamina Paton Fleming. *Astrophys. J.*, 1911, v.34, 314–317.

38. Spradley J.L. The Industrious Mrs. Fleming. *Astronomy*, 1990, v.18(7), 48–51.

39. Spradley J.L. Two centennials of star catalogs compiled by women. *Astron. Quart.*, 1990, v.7(3), 177–184.

40. Lankford J. and Slavings R.L. Gender and Science: Women in American Astronomy, 1859–1940. *Phys. Today*, 1990, 43(3), 58–65.

LETTERS TO PROGRESS IN PHYSICS

Pierre-Marie Luc Robitaille: A Jubilee Celebration

Dmitri Rabounski
E-mail: rabounski@ptep-online.com

We celebrate the 50th birthday anniversary of Prof. Pierre-Marie Robitaille, the author of Progress in Physics who is one of the leading experts in the Nuclear Magnetic Resonance Imaging. Prof. Robitaille is known as the designer of the most world's first Ultra High Field MRI scanner. Prof. Robitaille still continues his creative research activity in the field of thermal physics, connected to the origin of the Microwave Background and astrophysics.

July 12, 2010 marks the 50th birthday of Professor Pierre-Marie Robitaille. He was born in North Bay, Ontario, the third of ten children to Noel Antoine Robitaille and Jacqueline Alice Roy. Noel Robitaille had moved to Ontario from his native Quebec when he was stationed as a physician in the Royal Canadian Air Force. Eventually settling in northern Ontario, he served the villages of Massey and Espanola. In his role as a local doctor, Noel Robitaille would also care for the Ojibway population of the region. In 1964, he would be honored by the Ojibway Nation, becoming the first white man to bear the distinction of Ojibway chief of the Spanish River Band. His Indian name, *Ke-chutwa-ghizhigud*, meaning "Chief Holiday" [1].

Raised by French-Canadian parents, Pierre-Marie Robitaille attended L'École St. Joseph in Espanola, Ontario, where he studied primarily in his native tongue. Upon completion of the 8th grade, he attended Espanola High School, where education was conducted in English. As an adolescent, he often served as an altar boy during daily mass at St. Louis de France Catholic Church, the French parish of his community. Surrounded by the forests of Northern Ontario, he enjoyed ice fishing, hunting, and building log cabins in the woods.

In 1978, just as Robitaille was completing his secondary education, his father relocated to Cedar Falls, Iowa. Mrs. Robitaille and her children were to remain in northern Ontario. In order to maintain ties with his father, Robitaille enrolled at the University of Iowa in Iowa City. It was there that he met his future wife, Patricia. Though he relocated to Iowa for the 1978–1979 school year, Robitaille rarely saw his father. Therefore, he moved to Cedar Falls, Iowa. He would graduate from the University of Northern Iowa, in 1981, with a degree in general science.

At that time, Robitaille entered a Ph.D. program in biochemistry under the tutelage of Dr. David E. Metzler at Iowa State University, obtaining an M.S. degree in 1984. His masters thesis involved NMR equilibrium analysis of polyamines with vitamin B6. At the same time, Robitaille realized that in-vivo NMR was beginning to grow. He sought unsuccessfully to convince Dr. Metzler to enter this promising new area of biochemistry and, eventually, entered the field on his own.

Prof. Pierre-Marie Robitaille.

He transferred his graduate appointment to the Department of Zoology, where he brought in-vivo NMR methods to the laboratory of George Brown, an electron microscopist. It was there that he acquired a set of standards for in-vivo ^{31}P-NMR [2] and conducted some of the first studies of isolated sperm cells with ^{31}P-NMR [3, 4]. At the same time, Robitaille enrolled in the Inorganic Chemistry doctoral program, under the guidance of Professor Donald Kurtz. He graduated from Iowa State University with a Ph.D. in 1986, holding majors in Zoology and Inorganic Chemistry. His dissertation was divided into two parts which he would defend in front of separate committees, one for each major.

Following his Ph.D. training, Pierre-Marie Robitaille joined the in-vivo NMR laboratory of Professor Kamil Ugurbil at the University of Minnesota. There, he conducted work in cardiac spectroscopy, operating one of the first small animal 4.7T/40cm magnetic resonance instruments in the United

States. It was Professor Ugurbil who urged Robitaille to apply for faculty positions in magnetic resonance imaging and spectroscopy. Ultimately, he accepted the position of Director of Magnetic Resonance Research and Assistant Professor of Radiology at The Ohio State University, with a startup package well in excess of $1 million. He was 28.

While at Ohio State, Professor Robitaille established himself as a leader in cardiac spectroscopy and magnetic resonance [5, 6]. He would eventually design and assemble the world's first Ultra High Field MRI instrument [7–16]. The results obtained from this scanner would propel MRI into a new era in imaging technology. Professor Allan Elster, the Editor of the Journal of Computer Assisted Tomography recognized the magnitude of the contribution and arranged for a special issue of the journal to be published outlining some of the first 8 Tesla results. In is editorial comments relative to this issue, Dr. Elster wrote:

> "This is a landmark issue of the Journal of Computer Assisted Tomography. Contained within its pages are amazing images and technical descriptions of the world's first whole body human clinical magnetic resonance scanner operating at 8 Tesla. Congratulations to Pierre-Marie Robitaille and his co-workers in Radiology and Engineering at The Ohio State University for constructing a device some experts said would be impossible to build. The total stored magnetic energy in this 30,000 kg magnet is a remarkable 81 megajoules. To put this value into perspective, 81 MJoules is the kinetic energy of a 200-metric ton locomotive barreling down the track at 100 km per hour! The human images obtained so far are also astounding (Fig. 1), especially considering that the system has only been operational for a few months and many radio frequency coil and pulse sequence issues remain to be worked out. The Ohio State team has proposed a number of interesting theories concerning susceptibility effects and dielectric resonance phenomena within the human head at 8 Tesla. Some of these theories challenge traditional tenets in MR physics and are admittedly controversial. As more measurements are obtained and experiments are conducted, these theories will be refine, improved, or discarded. Robitaille et. al. have led us to a new frontier in clinical MR imaging. Perhaps one day in the not-so-distant future, 1.5 Tesla will be considered low-field imaging" [14].

The next month, Professor Robitaille established a new record for high resolution imaging in MRI, once again published in JCAT, with the following editorial note:

> "Pierre-Marie Robitaille and the Ohio State University MRI Team have done it again! In this issue they present the world's first MR images obtained at 2,000×2,000 resolution — in honor of the new millennium of course. In case you missed it, please check out the Journal of Computer Assisted Tomography's November/December 1999 issue. Here Robitaille and colleagues have published 10 landmark articles describing the design and construction of their 8 Tesla whole-body MR scanner, as well as additional remarkable images of the brain. If you wish to download some of the images directly (they look even better on a video monitor), please see the JCAT website at www.rad.bgsm.edu/jcat/supp.htm. Happy Y2K from all of us at JCAT!" [15].

The birth of Ultra High Field MRI represented a paradigm shift for many in the MRI community who had previously believed that human images could never be acquired at such field strengths [16, 17]. Relative to the creation of the first UHFMRI systems, Paul Lauterbur (Nobel Prize in Medicine and Physiology, 2003) wrote:

> "In the early machines, low radiofrequencies of 4 MHz or so meant that RF coil designs were simple (even inexperienced undergraduates could design and build such circuits with little knowledge of more than DC electrical circuits), and the forces on gradient coils were small. The effects of magnetic susceptibility inhomogeneity in and around the object being imaged were negligible, and RF penetration depths were not a problem for human-scale samples. Everything began to change as higher fields and higher frequencies came into use, and the earlier idyllic simplicities began to seem quaint. The trend continued, however, driven by the increased signal to noise ratios and the resultant higher resolution and speed available, and sophisticated engineering became more and more essential, not only for magnets but for gradient systems and radiofrequency transmitters and receivers, but also for better software for modeling and correcting distortions. Experts who had said, and even written, that frequencies above 10 MHz would never be practical watched in amazement as scientists and engineers pushed instrument performances to ever-higher levels at ever-increasing magnetic field strengths, as this volume demonstrates" [18].

Prior to assembling the 8 Tesla instrument, Professor Robitaille envisioned that his career would remain firmly grounded in MRI. However, the first results at 8 Tesla relative to RF power requirements in MRI profoundly altered his scientific outlook. He began to think about MRI as a thermal process. In the early days of NMR, the T1 relaxation time was referred to as the "thermal relaxation time". As a result, Professor Robitaille advanced the idea that, if MRI was thermal process, it should be possible to extract the temperature of the human head using the laws of thermal emission, in the same manner that Penzias and Wilson had measured a temperature of ∼3 K for the microwave background [19]. Unfortunately, such an approach yielded a Wien's displacement temperature of less than 1 K for the human head. Surely, something was

incorrect.

Professor Robitaille viewed magnetic resonance as enabling scientists to examine the reverse of the emission problem in the infrared, as studied by Planck and his predecessors [20]. Therefore, he turned his attention to thermal radiation and astrophysics. Soon, he published an abstract which questioned the assignment of the microwave background to the cosmos [21]. Then, in a bold step, he placed an ad in the New York Times [22] announcing the Collapse of the Big Bang and the Gaseous Sun. The response from the popular press and the scientific community was immediate and sometimes harsh [23–25]. Despite claims to the contrary [23], Professor Robitaille's advertisement in the New York Times had nothing to do with the concurrent debate in Ohio relative to evolution [23]. The timing was purely coincidental.

Following the ad in the New York Times, Professor Robitaille turned to Progress in Physics and began outlining his ideas in a series of papers which spanned a very broad area of fundamental physics. His papers on the WMAP [26] and COBE [27] satellites are amongst the most viewed by the journal audience and, eventually, his position was found to merit some consideration by the astrophysics community [28].

The study of Kirchhoff's Law of Thermal Emission has been the driving force behind Prof. Robitaille's work in astrophysics. Robitaille has demonstrated the invalidity of this law and its subsequent claims for universality [29–33]. Prof. Robitaille has also argued that the proper analysis of thermal emission should be attributed to Balfour Stewart [32]. Resting on the knowledge that Kirchhoff's Law was invalid, Robitaille argued for a liquid model of the Sun [34] and advanced simple proofs to strengthen his position [35]. Robitaille maintains that the emission of a thermal spectrum from the Sun, by itself, comprises all the proof necessary for a liquid model. Given the error within Kirchhoff Law, the Sun cannot be a gaseous plasma. It must be condensed matter.

Robitaille has also based his re-assignment of the microwave background to the Earth on Kirchhoff's Law [28]. He has shown that astrophysics did not properly consider the emission of water itself when contemplating the background [36, 37]. His recent paper analyzing the Planck satellite [38] further builds on his position, along with papers by the authors, Rabounski and Borissova [39]. Finally, Robitaille has questioned the validity of Boltzmann's constant [40]. This is the result of the correction of Kirchhoff's Law [28, 29–33] and the re-assignment of the microwave background to the Earth [36–39].

Robitaille maintains a quite lifestyle in Columbus, Ohio. He has been married to Patricia for 30 years, and they have three sons: Jacob, Christophe, and Luc. Dr. Robitaille enjoys sailing his Flying Scot and is an avid builder of timberframe structures.

The authors would like to acknowledge the contributions of Patricia Anne Robitaille in providing some factual details and photographs of her husband.

Submitted on July 12, 2011 / Accepted on July 13, 2011

References

1. Dr. N. Robitaille made indian chief. *Sudbury Star*, February 6, 1964.
2. Robitaille P.-M.L., Robitaille P.A., Brown G.G., Jr., and Brown G.G. An analysis of the pH-dependent chemical shift behavior of phosphorus containing metabolites. *J. Magn. Reson.*, 1990, v.92, 73–84.
3. Robitaille P.A., Robitaille P.M.L., and Brown G.G. ^{31}P NMR studies of Limulus Polyphemus: spermatozoa at rest and after motility. *J. Exp. Zool.*, 1986, v.238, 89–98.
4. Robitaille P.M.L., Robitaille P.A., Martin P.A., and Brown G.G. ^{31}P nuclear magnetic resonance studies of spermotozoa from the boar, ram, goat and bull. *Comp. Biochem. Physiol. B*, 1987, v.87(2), 285–296.
5. Berliner L.J. and Robitaille P-M.L. (Eds.) Biological magnetic resonance. Volume 15: In-vivo Carbon-13 NMR. Kluwer Academic and Plenum Publishers, New York, 1998.
6. Irsik R.D., White R.D., and Robitaille P-M.L. "Cardiac magnetic resonance imaging" in Moss and Adams heart disease in infants, children, and adolescents. 5th Edition, G. C. Emmanouilides, H. D. Allen, T. A. Riemenschneider, and H. P. Gutgesell, eds., Williams and Wilkins, Baltimore, 1995, v.1, 206–223.
7. 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.
8. Robitaille P.-M.L. and Berliner L.J. (Eds.) Biological magnetic resonance. Volume 26: Ultra high field magnetic resonance imaging. Springer, New York, 2006.
9. Robitaille P.-M.L. Magnetic resonance imaging and spectroscopy at very high fields: a step towards 8 Tesla. *Proceedings of Physical Phenonema at High Magnetic Fields-III*, Zachary Fisk, Lev Gor'kov, David Meltzer and Robert Schrieffer, eds., World Scientific, London, 1999, 421–426.
10. Robitaille P.-M.L. Magnetic resonance imaging. *McGraw-Hill 2000 Yearbook of Science and Technology*, McGraw-Hill Publishing Company, New York, N.Y., 2000, 235–237.
11. Steinberg D. Brain imaging assumes greater power, precision. *The Scientist*, 1998, v.12(8), 1.
12. Valeo T. Searching for a powerful sharper image: powerful MRI can spot tumors, aneurysms and mini-strokes, but can't yet reveal if the author's brain will bleed again. *Neurology Now*, 2006, v.2(3), 18–21.
13. Mahoney D. Delivering results: a progress report on brain research. DANA Foundation, 2002. Available online http://www.dana.org/uploadedFiles/The_Dana_Alliances/European_Dana_Alliance_for_the_Brain/progress-report-2000_en.pdf
14. Elster A.D. Clinical MR imaging at 8 Tesla: the new frontier of ultra high field MRI. *J. Comp. Assist. Tomogr.*, 1999, v.23(6), 807.
15. Elster A.D. JCAT @ Y2K. *J. Comp. Assist. Tomogr.*, 2000, v.24(1) , 1.
16. Robitaille P.-M.L. Response to "Does RF brain heating decrease at 8 T". *NMR Biomed.*, 1999, v.12, 256.
17. Robitaille P.-M.L. On RF power and dielectric resonances in UHFMRI. *NMR Biomed.*, 1999, v.12, 318–319.
18. Lauterbur P. Preface in Biological Magnetic Resonance, vol.26: Ultra high field magnetic resonance imaging. P.-M. Robitaille and L. J. Berliner, eds., Springer, New York, 2006, xiii.
19. Penzias A.A. and Wilson R.W. A measurement of excess antenna temperature at 4080 Mc/s. *Astrophys. J.*, 1965, v.1, 419–421.

20. Robitaille P.-M.L. The reverse of the Planckian experiment. *American Physical Society 2004 March Meeting*, 2004, March 22–26, Montreal, Canada, Y35.012.
21. Robitaille P.-M.L. Nuclear magnetic resonance and the age of the Universe. *American Physical Society Centennial Meeting*, 1999, March 19–26, Atlanta, Georgia, 1999, BC19.14.
22. Robitaille P.-M.L. The collapse of the Big Bang and the gaseous Sun. *New York Times*, March 17, 2002. Accessed online from http://www.thermalphysics.org/pdf/times.pdf
23. Ganz J. Ripples in Ohio from Ad on the Big Bang. *New York Times*, March 19, 2002.
24. Everything has its price... *European Research News Center*, October 16, 2002. Available online http://ec.europa.eu/research/news-centre/en/soc/02-09-special-soc02.html
25. Park B. Alternative publishing: communicating science by full-page ad. *What's New by Bob Park*, March 29, 2002. Available online http://bobpark.physics.umd.edu/WN02/wn032902.html
26. Robitaille P.M. WMAP: a radiological analysis. *Progr. in Phys.*, 2007, v.1, 3–18.
27. Robitaille P.M. COBE: a radiological analysis. *Progr. in Phys.*, 2009, v.4, 17–42.
28. Robitaille P.-M.L. A radically different point of view on the CMB. *Questions of Modern Cosmology*, Mauro D'Onofrio and Carlo Burigana, eds., Springer Scientific, New York, N.Y., 2009, 93–108.
29. Robitaille P.M.L. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v.31(6), 1263–1267.
30. Robitaille P.M.L. An analysis of universality in blackbody radiation. *Prog. in Phys.*, 2006, v.2, 22–23.
31. Robitaille P.M.L. Blackbody radiation and the carbon particle. *Prog. in Phys.*, 2008, v.3, 36–55.
32. Robitaille P.M.L. A critical analysis of universality and Kirchhoff's law: a return to Stewart's law of thermal emission. *Prog. in Phys.*, 2008, v.3, 30–35.
33. Robitaille P.M.L. Kirchhoff's Law of Thermal Emission: 150 years. *Prog. in Phys.*, 2009, v.4, 3–13.
34. Robitaille P.M.L. A high temperature liquid plasma model of the Sun. *Prog. in Phys.*, 2007, v.1, 71–81.
35. Robitaille P.M.L. The solar photosphere: evidence for condensed matter. *Prog. in Phys.*, 2006, v.2, 17–21.
36. Robitaille P.M.L. The Earth Microwave Background (EMB), atmospheric scattering and the generation of isotropy. *Prog. in Phys.*, 2008, v.2, 164–165.
37. Robitaille P.M.L. Water, hydrogen bonding, and the microwave background. *Progr. in Phys.*, 2009, v.2, L5–L8.
38. Robitaille P.M.L. The Planck satellite LFI and the Microwave Background: importance of the 4 K reference targets. *Prog. in Phys.*, 2010, v.3, 11–18.
39. Borissova L. and Rabounski D. PLANCK, the satellite: a new experimental test of General Relativity. *Prog. in Phys.*, 2008, v.2, 3–14.
40. Robitaille P.M.L. Blackbody radiation and the loss of universality: implications for Planck's formulation and Boltzman's constant. *Prog. in Phys.*, 2009, v.4, 14–16.

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), Zentralblatt MATH (Germany), 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,
200 College Road, Gallup, NM 87301, USA

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

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

October 2011 **VOLUME 4**

CONTENTS

Information for Authors and Subscribers

Progress in Physics has been created for 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. 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. We accept brief papers, no larger than 8 typeset journal pages. Short articles are preferable. Large papers can be considered in exceptional cases to the section *Special Reports* intended for such publications in the journal. 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.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

Ultracold Fermi and Bose Gases and Spinless Bose Charged Sound Particles

Vahan N. Minasyan and Valentin N. Samoylov

Scientific Center of Applied Research, JINR, Dubna, 141980, Russia.
E-mail: mvahan_n@yahoo.com

We propose a novel approach for investigation of the motion of Bose or Fermi liquid (or gas) which consists of decoupled electrons and ions in the uppermost hyperfine state. Hence, we use such a concept as the fluctuation motion of "charged fluid particles" or "charged fluid points" representing a charged longitudinal elastic wave. In turn, this elastic wave is quantized by spinless longitudinal Bose *charged sound particles* with the rest mass m and charge e_0. The existence of spinless Bose *charged sound particles* allows us to present a new model for description of Bose or Fermi liquid via a non-ideal Bose gas of *charged sound particles*. In this respect, we introduce a new postulation for the superfluid component of Bose or Fermi liquid determined by means of *charged sound particles* in the condensate, which may explain the results of experiments connected with ultra-cold Fermi gases of spin-polarized hydrogen, ^{6}Li and ^{40}K, and such a Bose gas as ^{87}Rb in the uppermost hyperfine state, where the Bose-Einstein condensation of *charged sound particles* is realized by tuning the magnetic field.

1 Introduction

The Bose-Einstein condensation (BEC) has a wide application for investigation of superconductivity of metals and superfluidity of liquids. The primary experimental challenge to evaporative cooling of spin-polarized hydrogen was made by a dilution refrigerator, demonstrating that spin-polarized hydrogen can be confined in a statistic magnetic trap and thermally decoupled from the walls [1–3]. At the density $\frac{N}{V} \approx 10^{13}$ cm^{-3} it is observed that the gas consisting of decoupled electrons and ions in the uppermost hyperfine state is evaporatively cooled to a temperature approximately equal to 40 mK.

Here, we remark about BEC that was produced in a vapor of ^{87}Rb bosonic ions confined by magnetic fields and evaporatively cooled [4]. The condensate fraction first appeared near a temperature of 170 nanokelvin at the density $\frac{N}{V} = 2.6 \times 10^{12}$ cm^{-3}. The experiment has shown that the value of temperature 170 nK is reduced to 20 nK. In reality, the strongly interacting spin-$\frac{1}{2}$ ^{6}Li and ^{40}K fermionic gases were realized via tuning the magnetic field [5]. These experimental achievements in the field of ultra-cold Fermi gases are based mainly on the possibility of tuning the scattering length a which becomes much larger in magnitude than the mean interatomic distance by changing the external magnetic field. In this respect, the concept of Fermi surface loses its meaning due to the broadening produced by pairing of fermions, the so-called Feshbach resonances in ultracold atomic Fermi gases. However, in this letter we predict a new method of liquid cooling which is based on the formation of oscillators at every point of liquid by tuning the magnetic field, which in turn leads to vibration of "charged fluid particles". These "charged fluid particles" reproduce charged spinless quasiparticles which determine the superfluidity component of Bose or Fermi liquid by action of the static magnetic field.

In order to investigate the motion of quantum liquid (or quantum gas) in the uppermost hyperfine state, we consider the motion of "charged fluid particles" by means of a charged longitudinal elastic wave [6]. This longitudinal elastic wave is quantized by spinless Bose *charged sound particles* with the mass m and charge e_0. Further, we present a new model for description of charged Bose or Fermi liquid via a non-ideal Bose gas consisting of *charged sound particles*. As opposed to London's postulation about the superfluid component of liquid ^{4}He [7], we introduce a new postulation about the superfluid component of Bose or Fermi liquid via *charged sound particles* in the condensate. On the other hand, we estimate the zero sound speed which leads to the correct explanation of the experimental result connected with the BEC of a gas consisting of spin-polarized hydrogen.

2 Quantization of quantum liquid or quantum gas in the uppermost hyperfine state

Now let us analyze quantization of quantum liquid (or quantum gas) in the uppermost hyperfine state. This quantum liquid (or quantum gas) consists of N Bose or Fermi positive charged ions with the charge e and mass M confined in the volume V where they are in a negative electron background since the entire system of liquid is electro-neutral. Considering quantum liquid as a continuous medium, we investigate the fluctuation motion of the number n of "charged fluid particles" on the basis of hydrodynamics (where a "charged fluid particle" is defined as a very small volume V_0 in regard to the volume V of the liquid ($V_0 \ll V$) with the mass m and charge e_0. The volume V_0 contains the number $N' = \frac{N}{n}$ of liquid ions, therefore the charge e_0 is expressed via the ion charge as $e_0 = \frac{eN}{n}$.

In accordance with the laws of hydrodynamics [6], the mass density ρ and pressure p of liquid are presented as

$$\rho = \rho_0 + \rho^{'}$$

and

$$p = p_0 + p^{'},$$

where $\rho_0 = \frac{MN}{V}$ and p_0 are, respectively, the equilibrium mass density and pressure; $\rho^{'}$ and $p^{'}$ are the relative fluctuations of the mass density and pressure.

As is known, the continuity equation has the form:

$$\frac{\partial \rho^{'}}{\partial t} = -\rho_0 \operatorname{div} \vec{v}, \tag{1}$$

which may present as:

$$\rho^{'} = -\rho_0 \operatorname{div} \vec{u}, \tag{2}$$

where $\vec{v} = \frac{\partial \vec{u}}{\partial t}$ is the speed of a charged fluid particle; $\vec{u} = \vec{u}(\vec{r}, t)$ is the displacement vector of a charged fluid particle which describes a charged longitudinal sound wave.

On the other hand, Euler's equation in the first-order-of-smallness approximation takes the reduced form:

$$\frac{\partial \vec{v}}{\partial t} + \frac{\nabla p^{'}}{\rho_0} = 0. \tag{3}$$

Hence, we consider the fluctuation motion of charged fluid particles as adiabatic, deriving the following equation:

$$p^{'} = \left(\frac{\partial p}{\partial \rho_0}\right)_S \rho^{'} = c_l^2 \rho^{'}, \tag{4}$$

where S is the entropy; $c_l = \sqrt{\left(\frac{\partial p}{\partial \rho_0}\right)_S}$ is the speed of the charged longitudinal elastic wave.

As is known, the fluctuation motion of charged fluid particles represents as a potential one:

$$\operatorname{curl} \vec{v} = \operatorname{curl} \frac{\partial \vec{u}}{\partial t} = 0. \tag{5}$$

Thus, by using the above equation we may get to the wave equation for the vector of displacement $\vec{u} = \vec{u}(\vec{r}, t)$:

$$\nabla^2 \vec{u}(\vec{r}, t) - \frac{1}{c_l^2} \frac{\partial^2 \vec{u}(\vec{r}, t)}{\partial t^2} = 0, \tag{6}$$

which in turn describes the longitudinal charged sound wave.

Now, we state that the longitudinal elastic wave consists of spinless Bose *charged sound particles* with the non-zero rest mass m. Then, the displacement vector $u(\vec{r}, t)$ is expressed via a secondary quantization vector of the wave function of spinless Bose *charged sound particles* directed along the wave vector $\vec{k}$:

$$\vec{u}(\vec{r}, t) = u_l \left(\vec{\phi}(\vec{r}, t) + \vec{\phi}^{+}(\vec{r}, t)\right), \tag{7}$$

where u_l is the normalization constant which is the amplitude of oscillations; $\vec{\phi}(\vec{r}, t)$ is the secondary quantization of vector wave functions for creation and annihilation of one longitudinal *charged sound particle* with the mass m whose direction $\vec{l}$ is directed towards the wave vector $\vec{k}$:

$$\vec{\phi}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \vec{a}_{\vec{k}}\, e^{i(\vec{k}\vec{r} - kc_l t)} \tag{8}$$

$$\vec{\phi}^{+}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \vec{a}^{+}_{\vec{k}}\, e^{-i(\vec{k}\vec{r} - kc_l t)} \tag{9}$$

with the condition

$$\int \vec{\phi}^{+}(\vec{r}, t)\, \vec{\phi}(\vec{r}, t)\, dV = n_0 + \sum_{\vec{k} \neq 0} \hat{a}^{+}_{\vec{k}}\, \hat{a}_{\vec{k}} = \hat{n}, \tag{10}$$

where $\vec{a}^{+}_{\vec{k}}$ and $\vec{a}_{\vec{k}}$ are, respectively, the Bose vector-operators of creation and annihilation for a free *charged sound particle* with the energy $\frac{\hbar^2 k^2}{2m}$, described by the vector $\vec{k}$ whose direction coincides with the direction $\vec{l}$ of a traveling charged longitudinal elastic wave; $\hat{n}$ is the operator of the total number of *charged sound particles*; $\hat{n}_0$ is the total number of *charged sound particles* at the condensate level with the wave vector $\vec{k} = 0$.

Thus, as is seen, the displacement vector $\vec{u}(\vec{r}, t)$ satisfies wave-equation (6) and in turn takes the form:

$$\vec{u}(\vec{r}, t) = \vec{u_0} + \frac{u_l}{\sqrt{V}} \sum_{\vec{k} \neq 0} \left(\vec{a}_{\vec{k}}\, e^{i(\vec{k}\vec{r} - kc_l t)} + \vec{a}^{+}_{\vec{k}}\, e^{-i(\vec{k}\vec{r} - kc_l t)}\right). \tag{11}$$

While investigating a superfluid liquid, Bogoliubov [8] separated the atoms of helium in the condensate from those atoms filling the states above the condensate. In an analogous manner, we may consider the vector operator $\vec{a}_0 = \vec{l}\sqrt{n_0}$ and $\vec{a}^{+}_0 = \vec{l}\sqrt{n_0}$ as c-numbers (where $\vec{l}$ is the unit vector in the direction of propagation of the sound wave) within the approximation of a macroscopic number of *sound particles* in the condensate $n_0 \gg 1$. These assumptions lead to a broken Bose-symmetry law for *sound particles* in the condensate. To extend the concept of a broken Bose-symmetry law for *sound particles* in the condensate, we apply the definition of BEC of *sound particles* in the condensate as was postulated by the Penrose-Onsager for the definition of BEC of helium atoms [9]:

$$\lim_{n_0, n \to \infty} \frac{n_0}{n} = const. \tag{12}$$

On the other hand, we may observe that presence of *charged sound particles* filling the condensate level with the wave vector $\vec{k} = 0$ leads to the appearance of the constant displacement $\vec{u}_0 = \frac{2u_l \vec{l}\sqrt{n_0}}{\sqrt{V}}$ of *charged sound particles*.

To find the normalization constant u_l, we introduce the following condition which allows us to suggest that at absolute zero all *sound particles* fill the condensate level $\vec{k} = 0$.

This reasoning implies that at $n_0 = n$ the constant displacement takes the maximal value $2d = \sqrt{|\vec{u}_0|^2}$ which represents the maximal distance between two neighboring *charged sound particles*. On the other hand, this distance is determined by the formula $d = \left(\frac{3V}{4\pi n}\right)^{\frac{1}{3}}$, which is in turn substituted into the expression $2d = \sqrt{|\vec{u}_0|^2}$. Then, consequently, we get to the normalization constant $u_l = 0.65\left(\frac{n}{V}\right)^{-\frac{5}{6}}$.

The condition for conservation of density at each point of liquid stipulates that

$$\rho_0 = \frac{MN}{V} = \frac{mn}{V}, \tag{13}$$

which represents a connection of the mass and density of the *charged sound particles* with the mass and density of the ions. Thus, we argue that liquid (or gas) can be described by the model of an ideal gas of *n charged sound particles* with the mass m and charge e_0 in the volume V. Hence, we remark that the Coulomb scattering between *charged sound particles* is neglected in the considered theory.

3 "Charged fluid particles" in trapped static magnetic field

Now, we consider the Hamiltonian operator $\hat{H}_l$ of liquid [6] in a trapped static magnetic field [10]:

$$\hat{H}_l = \frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV + \frac{1}{2}\int\left(\frac{c_l\rho^{'}}{\sqrt{\rho_0}}\right)^2 dV + \\ + \frac{\rho_0}{2}\int(\Omega\vec{u}_l)^2\, dV, \tag{14}$$

where $\Omega = \frac{e_0 H}{mc}$ is the trapping frequency of a "charged fluid particle"; e_0 is the charge of a "fluid particle"; H is the absolute value of the magnetic strain; c is the velocity of light in vacuum. Hence, we note that the charge of a fluid particle equals $e_0 = eN^{'} = \frac{Ne}{n}$, where $N^{'}$ is the number of ions in a small volume V_0 of one charged fluid particle.

Substituting $\rho^{'}$ from (2) into (14), we obtain

$$\hat{H}_l = \frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV + \frac{\rho_0}{2}\int(c_l\,\mathrm{div}\,\vec{u})^2\, dV + \\ + \frac{\rho_0}{2}\int(\Omega\vec{u}_l)^2\, dV. \tag{15}$$

Using Dirac's approach in [11] for quantization of the electromagnetic field, we have:

$$\frac{\partial\vec{u}(\vec{r},t)}{\partial t} = -\frac{ic_l\vec{u}_l}{\sqrt{V}}\sum_{\vec{k}} k\left(\vec{a}_{\vec{k}}\, e^{-ikc_l t} - \vec{a}^{+}_{-\vec{k}}\, e^{ikc_l t}\right)e^{i\vec{k}\vec{r}}, \tag{16}$$

as well as

$$\mathrm{div}\,\vec{u}(\vec{r},t) = \frac{i\vec{u}_l}{\sqrt{V}}\sum_{\vec{k}}\vec{k}\left(\vec{a}_{\vec{k}}\, e^{-ikc_l t} + \vec{a}^{+}_{-\vec{k}}\, e^{ikc_l t}\right)e^{i\vec{k}\vec{r}}. \tag{17}$$

Now, introducing (16) and (17) into (15) and using

$$\frac{1}{V}\int e^{i(\vec{k}_1+\vec{k}_2)\vec{r}} = \delta^3_{\vec{k}_1+\vec{k}_2},$$

we obtain the terms in the right side of the Hamiltonian of the system presented in (15):

$$\frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV = -\frac{\rho_0 c_l^2 u_l^2}{2}\sum_{\vec{k}} k^2\left(\vec{a}_{\vec{k}} - \vec{a}^{+}_{-\vec{k}}\right)\left(\vec{a}_{-\vec{k}} - \vec{a}^{+}_{\vec{k}}\right),$$

$$\frac{\rho_0}{2}\int(\mathrm{div}\,\vec{u})^2\, dV = \frac{\rho_0 c_l^2 u_l^2}{2}\sum_{\vec{k}} k^2\left(\vec{a}_{\vec{k}} + \vec{a}^{+}_{-\vec{k}}\right)\left(\vec{a}_{-\vec{k}} + \vec{a}^{+}_{\vec{k}}\right)$$

and

$$\frac{\rho_0}{2}\int(\Omega\vec{u}_l)^2\, dV = \frac{\rho_0\Omega^2 u_l^2}{2}\sum_{\vec{k}}\left(\vec{a}_{\vec{k}} + \vec{a}^{+}_{-\vec{k}}\right)\left(\vec{a}_{-\vec{k}} + \vec{a}^{+}_{\vec{k}}\right).$$

These expressions determine the reduced form of the Hamiltonian operator $\hat{H}_l$ by the form:

$$\hat{H}_l = \sum_{\vec{k}}\left(2\rho_0 u_l^2 c_l^2 k^2 + \rho_0\Omega^2 u_l^2\right)\vec{a}^{+}_{\vec{k}}a_{\vec{k}} + \\ + \frac{\rho_0\Omega^2 u_l^2}{2}\sum_{\vec{k}}\left(\vec{a}^{+}_{-\vec{k}}\vec{a}^{+}_{\vec{k}} + \vec{a}_{\vec{k}}\vec{a}_{-\vec{k}}\right), \tag{18}$$

where u_l^2 is defined by the first term in the right side of (18) which represents the kinetic energy of a *charged sound particle* $\frac{\hbar^2k^2}{2m}$, if we suggest:

$$2\rho_0 u_l^2 c_l^2 k^2 = \frac{\hbar^2 k^2}{2m}. \tag{19}$$

Then,

$$u_l^2 = \frac{\hbar^2}{4c_l^2 m\rho_0},$$

which allows one to determine the mass m of a *charged sound particle* using the value of the normalization constant $u_l = 0.65\left(\frac{n}{V}\right)^{-\frac{5}{6}}$ and (13):

$$m = \frac{\hbar}{c_l}\left(\frac{n}{V}\right)^{\frac{1}{3}}. \tag{20}$$

Thus, the main part of the Hamiltonian operator $\hat{H}_l$ takes the form:

$$\hat{H}_l = \sum_{\vec{k}\neq 0}\left(\frac{\hbar^2k^2}{2m} + mv^2\right)\vec{a}^{+}_{\vec{k}}\,a_{\vec{k}} + \frac{mv^2}{2}\sum_{\vec{k}\neq 0}\left(\vec{a}^{+}_{-\vec{k}}\,\vec{a}^{+}_{\vec{k}} + \vec{a}_{\vec{k}}\,\vec{a}_{-\vec{k}}\right), \tag{21}$$

where we denote $v = \frac{\hbar\Omega}{\sqrt{2}mc_l}$, which in turn is the speed of charged sound in a Bose or Fermi liquid excited by static magnetic field; n_0 is the number of *charged sound particles* in the condensate.

For the evolution of the energy level, it is necessary to diagonalize the Hamiltonian $\hat{H}_l$, which can be accomplished by introducing the vector Bose-operators $\vec{b}^+_{\vec{k}}$ and $\vec{b}_{\vec{k}}$ [12]:

$$\vec{a}_{\vec{k}} = \frac{\vec{b}_{\vec{k}} + L_{\vec{k}}\vec{b}^+_{-\vec{k}}}{\sqrt{1 - L^2_{\vec{k}}}}, \tag{22}$$

where $L_{\vec{k}}$ is the unknown real symmetrical function of the wave vector $\vec{k}$.

By substituting (22) into (21), we obtain

$$\hat{H}_l = \sum_{\vec{k}\neq 0} \varepsilon_{\vec{k}} \vec{b}^+_{\vec{k}} \vec{b}_{\vec{k}}, \tag{23}$$

where $\vec{b}^+_{\vec{k}}$ and $\vec{b}_{\vec{k}}$ are the creation and annihilation operators of charged Bose quasiparticles with the energy:

$$\varepsilon_{\vec{k}} = \left[\left(\frac{\hbar^2 k^2}{2m}\right)^2 + \hbar^2 k^2 v^2\right]^{1/2}. \tag{24}$$

In this context, the real symmetrical function $L_{\vec{k}}$ of the wave vector $\vec{k}$ is found to be

$$L^2_{\vec{k}} = \frac{\frac{\hbar^2 k^2}{2m} + mv^2 - \varepsilon_{\vec{k}}}{\frac{\hbar^2 k^2}{2m} + mv^2 + \varepsilon_{\vec{k}}}. \tag{25}$$

Thus, the average energy of the system takes the form:

$$\overline{\hat{H}_l} = \sum_{\vec{k}\neq 0} \varepsilon_{\vec{k}} \overline{\vec{b}^+_{\vec{k}} \vec{b}_{\vec{k}}}, \tag{26}$$

where $\overline{\vec{b}^+_{\vec{k}} \vec{b}_{\vec{k}}}$ is the average number of charged Bose quasiparticles with the wave vector $\vec{k}$ at the temperature T:

$$\overline{\vec{b}^+_{\vec{k}} \vec{b}_{\vec{k}}} = \frac{1}{e^{\frac{\varepsilon_{\vec{k}}}{kT}} - 1}. \tag{27}$$

Thus, we have found the spectrum of free charged spinless quasiparticles excited in a Bose or Fermi liquid which is similar to Bogoliubov's one [8]. In fact, the Hamiltonian of system (24) describes an ideal Bose gas consisting of charged spinless phonons at a small wave number $k \ll \frac{2mv}{\hbar}$ but at $k \gg \frac{2mv}{\hbar}$ the Hamiltonian operator describes an ideal gas of *charged sound particles*. This reasoning implies that the tuning magnetic field forms the superfluidity component of a Bose or Fermi liquid which is been in the uppermost hyperfine state.

4 BEC of *charged sound particles*

As opposed to London's postulation concerning BEC of atoms [7], we state that *charged sound particles* in the condensate define the superfluid component of Bose and Fermi liquids. Consequently, statistical equilibrium equation (10) takes the following form:

$$n_{0,T} + \sum_{\vec{k}\neq 0} \overline{\vec{a}^+_{\vec{k}} \vec{a}_{\vec{k}}} = n, \tag{28}$$

where $\overline{\vec{a}^+_{\vec{k}} \vec{a}_{\vec{k}}}$ is the average number of *charged sound particles* with the wave vector $\vec{k}$ at the temperature T.

To find the form $\overline{\vec{a}^+_{\vec{k}} \vec{a}_{\vec{k}}}$, we use the linear transformation presented in (22):

$$\overline{\vec{a}^+_{\vec{k}} \vec{a}_{\vec{k}}} = \frac{1 + L^2_{\vec{p}}}{1 - L^2_{\vec{p}}} \overline{\vec{b}^+_{\vec{k}} \vec{b}_{\vec{k}}} + \frac{L_{\vec{k}}}{1 - L^2_{\vec{k}}} \left(\overline{\vec{b}^+_{\vec{k}} \vec{b}^+_{-\vec{k}}} + \overline{\vec{b}_{\vec{k}} \vec{b}_{-\vec{k}}}\right) + \frac{L^2_{\vec{k}}}{1 - L^2_{\vec{k}}}.$$

According to the Bloch-De-Dominicis theorem, we have

$$\overline{\vec{b}^+_{\vec{k}} \vec{b}^+_{-\vec{k}}} = \overline{\vec{b}_{\vec{k}} \vec{b}_{-\vec{k}}} = 0.$$

In this respect, the equation for the density of *charged sound particles* in the condensate takes the following form:

$$\frac{n_{0,T}}{V} = \frac{n}{V} - \frac{1}{V} \sum_{\vec{k}\neq 0} \frac{L^2_{\vec{k}}}{1 - L^2_{\vec{k}}} - \frac{1}{V} \sum_{\vec{k}\neq 0} \frac{1 + L^2_{\vec{k}}}{1 - L^2_{\vec{k}}} \overline{\vec{b}^+_{\vec{k}} \vec{b}_{\vec{k}}}. \tag{29}$$

Obviously, at the lambda transition $T = T_\lambda$ the density of *charged sound particles* $\frac{n_{0,T_\lambda}}{V} = 0$. Hence, we note that the mass m and density $\frac{n}{V}$ of *charged sound particles* are expressed via the mass of ions M and density of ions $\frac{N}{V}$ when solving a system of two equations presented in (13) and (20):

$$\frac{n}{V} = \left(\frac{Mc_l}{\hbar} \frac{N}{V}\right)^{\frac{3}{4}} \tag{30}$$

and

$$m = \left(\frac{\hbar}{c_l}\right)^{\frac{3}{4}} \left(\frac{MN}{V}\right)^{\frac{1}{4}}. \tag{31}$$

In conclusion, it should be noted that the given approach opens up a new direction for investigation of BEC of *charged sound particles* in Fermi gases of spin-polarized hydrogen, ^{6}Li and ^{40}K, and in a Bose gas such as ^{87}Rb, because the model of quantum liquid in the uppermost hyperfine state is considered in the same way as superfluid liquid helium. In this letter, we argue for the first time that the superfluid component of Bose or Fermi liquid in the uppermost hyperfine state is determined by means of *charged sound particles* in the condensate. In fact, we argue that the lambda transition point depends on the strain of static magnetic field due to equation (29) and condition for the density of *charged sound particles* $\frac{n_{0,T_\lambda}}{V} = 0$.

Submitted on April 14, 2011 / Accepted on April 18, 2011

References

1. Hess H. F. et al. Evaporative cooling of magneticaly trapped and compressed pin-olarized hydrogen. *Physical Review B*, 1986, v. 34, 3476–3479.
2. Hess H. F. et al. Magnetic Trapping of Spin-Polarized Atomic Hydrogen. *Physical Review Letters*, 1987, v. 59, 672–675.
3. Masuhara N. et al. Evaporative Cooling of Spin-Polarized Atomic Hydrogen. *Physical Review Letters*, 1988, v. 61, 935–938.
4. Anderson M. H. et al. Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor. *Science*, 1995, v. 269, 198.
5. Giorgini S., Pitaevskii L. P., Stringari S. Theory of ultracold atomic Fermi gases. *Reviews of Modern Physics*, 2008, v. 80, 1215–1274.
6. Landau L. D., Lifshitz E.,M. Theory of Elasticity. *Theoretical Physics*, 1986, v. 1, p. 350.
7. London F. The λ-Phenomenon of Liquid Helium and the Bose-Einstein Degeneracy. *Nature*, 1938, v. 141, 643–644.
8. Bogoliubov N. N. On the theory of superfludity. *Journal of Physics (USSR)*, 1974, v. 11, 23–32.
9. Penrose O., Onsager L. Bose-Einstein condensation and liquid helium. *Physical Review*, 1956, v. 104, 576–584.
10. Xia-Ji Liu et al. Virial Expansion for a Strongly Correlated Fermi Gas. *Physical Review Letters*, 2009, v. 102, 160401.
11. Dirac P. A. M. The Principles of Quantum Mechanics. Clarendon press, Oxford, 1958.
12. Minasyan V. N., Samoilov V. N. Two Type Surface Polaritons Excited into Nanoholes in Metal Films. *Progress in Physics*, 2010, v. 2, 3–6.

Superfluidity Component of Solid ^{4}He and *Sound Particles* with Spin 1

Vahan N. Minasyan and Valentin N. Samoylov

Scientific Center of Applied Research, JINR, Dubna, 141980, Russia.
E-mail: mvahan_n@yahoo.com

We present a new model for solid which is based on such a concept as the fluctuation motion of "solid particles" or "solid points". The fluctuation motion of "solid particles" in solid ^{4}He represents a longitudinal elastic wave which is in turn quantized by neutral longitudinal Bose *sound particles* with spin 1 with the rest mass m. Thus, first we remove a concept of "lattice" for solid by presentation of new model of one as a vibration of *sound particles* by natural frequency Ω_l. In this respect, we first postulate that the superfluid component of a solid ^{4}He is determined by means of *sound particles* with spin 1 in the condensate.

1 Introduction

The quantum solid is remarkable object which reveal macroscopic quantum phenomena, such as superfluidity and Bose-Einstein condensation (BEC) of solid ^{4}He [1] which were reported by many authors [2, 3].

The original theory proposed by Einstein in 1907 was of great historical relevance [4]. In the Einstein model, each atom oscillates relatively to its neighbors in the lattice which execute harmonic motions around fixed positions, the knots of the lattice. He treated the thermal property of the vibration of a lattice of N atoms as a $3N$ harmonic independent oscillator by identical own frequency Ω_0 which was quantized by application of the prescription developed by Plank in connection with the theory of Black Body radiation. The Einstein model could obtain the Dulong and Petit prediction at high temperature but could not reproduce an adequate representation of the the lattice at low temperatures. In 1912, Debye proposed to consider the model of the solid [5], by suggestion that the frequencies of the $3N$ harmonic independent oscillators are not equal as it was suggested by the Einstein model. In addition to his suggestion, the acoustic spectrum of solid may be treated as if the solid represented a homogeneous medium, except that the total number of independent elastic waves is cut off at $3N$, to agree with the number of degrees of freedom of N atoms. In this respect, Debye stated that one longitudinal and two transverse waves are excited in solid. These velocities of sound cannot be observed in a solid at frequencies above the cut-off frequency. Also, he suggested that phonon is a spinless. Thus, the Debye model correctly showed that the heat capacity is proportional to the T^3 law at low temperatures. At high temperatures, he obtained the Dulong-Petit prediction compatible to experimental results.

The other model of solid was presented by the authors of this letter in [6] where the solid was considered as continuum elastic medium consisting of neutral Fermi-atoms, fixed in the knots of lattice. In this case, we predicted that the lattice represents as the Bose-gas of Sound-Particles with finite masses m_l and m_t, corresponding to a longitudinal and a transverse elastic field. On the other hand, the lattice was considered as a new substance of matter consisting of sound particles, which excite the one longitudinal and one transverse elastic waves (this approach is differ from Debye one). These waves act on the Fermi-atoms which are stimulating a vibrations with the natural frequencies Ω_l and Ω_t. In this context, we introduced a new principle of elastic wave-particle duality, which allows us to build the lattice model. The given model leads to the same results as presented by Debye's theory.

However, we consider the model of solid by new way by introducing of such a concept as the fluctuation motion of "solid particles" or "solid points". In this respect, we remove a concept as a lattice of solid or an atoms, fixed in the knots of lattice because we deal with the "solid particle" which exist in any point of the solid. This "solid particle" is a similar to the "fluid particle" on the basis of hydrodynamics [7] (where "fluid particle" is determined as a very small volume V_0, in regard to the volume V of the liquid ($V_0 \ll V$), which consists of a macroscopic number of liquid atoms). The motion of "solid particle" describes the longitudinal elastic wave which in turn represents a Bose gas of neutral *sound particles* with spin 1 with finite mass m. In this letter, we present a new model of solid which describes a vibration of *sound particles* by natural frequency Ω_l. We postulate also that the superfluid component of a solid is determined by means of *sound particles* in the condensate.

2 Analysis

For beginning let us analyze quantization of a quantum liquid (or quantum gas) which consists of N Bose or Fermi atoms with the mass M confined in the volume V. Considering a quantum liquid as a continuum medium, we investigate the fluctuation motion of "fluid particles" on the basis of hydrodynamics (where "fluid particle" is determined as a very small volume V_0, in regard to the volume V of the liquid ($V_0 \ll V$), which consists of a macroscopic number of liquid atoms).

In accordance with the hydrodynamics laws, the mass

density ρ and pressure p for a liquid are presented as

$$\rho = \rho_0 + \rho^{'}$$

and

$$p = p_0 + p^{'},$$

where $\rho_0 = \frac{MN}{V}$ and p_0 are, respectively, the equilibrium mass density and pressure; $\rho^{'}$ and $p^{'}$ are the relative fluctuations of the mass density and pressure.

As is known, the continuity equation has the form:

$$\frac{\partial \rho^{'}}{\partial t} = -\rho_0 \operatorname{div} \vec{v}, \tag{1}$$

which may present as:

$$\rho^{'} = -\rho_0 \operatorname{div} \vec{u}, \tag{2}$$

where $\vec{v} = \frac{\partial \vec{u}}{\partial t}$ is the speed of a fluid particle; $\vec{u} = \vec{u}(\vec{r}, t)$ is the displacement vector of a fluid particle which describes a longitudinal sound wave.

On the other hand, Euler's equation in the first-order-of-smallness approximation takes the reduced form:

$$\frac{\partial \vec{v}}{\partial t} + \frac{\nabla p^{'}}{\rho_0} = 0. \tag{3}$$

Hence, we consider the fluctuation motion of fluid particles as adiabatic, deriving the following equation:

$$p^{'} = \left(\frac{\partial p}{\partial \rho_0}\right)_S \rho^{'} = c_l^2 \rho^{'}, \tag{4}$$

where S is the entropy of liquid; $c_l = \sqrt{\left(\frac{\partial p}{\partial \rho_0}\right)_S}$ is the speed of the longitudinal elastic wave.

As is known, the fluctuation motion of fluid particles represents as a potential one:

$$\operatorname{curl} \vec{v} = \operatorname{curl} \frac{\partial \vec{u}}{\partial t} = 0. \tag{5}$$

Thus, by using the above equation we may get to the wave equation for the vector of displacement $\vec{u} = \vec{u}(\vec{r}, t)$:

$$\nabla^2 \vec{u}(\vec{r}, t) - \frac{1}{c_l^2} \frac{\partial^2 \vec{u}(\vec{r}, t)}{\partial t^2} = 0, \tag{6}$$

which in turn gives a description of the longitudinal sound wave.

Now, we state that the longitudinal elastic wave consists of neutral spinless Bose *sound particles* with the non-zero rest mass m. Then, the displacement vector $u(\vec{r}, t)$ is expressed via a secondary quantization vector of the wave function of spinless Bose *sound particles* directed along the wave vector $\vec{k}$:

$$\vec{u}(\vec{r}, t) = u_l \left(\vec{\phi}(\vec{r}, t) + \vec{\phi}^{+}(\vec{r}, t)\right), \tag{7}$$

where u_l is the normalization constant which is the amplitude of oscillations; $\vec{\phi}(\vec{r}, t)$ is the second quantization vector wave functions for creation and annihilation of one longitudinal *sound particle* with the mass m whose direction $\vec{l}$ is directed towards the wave vector $\vec{k}$:

$$\vec{\phi}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \vec{a}_{\vec{k}}\, e^{i(\vec{k}\vec{r} - kc_l t)} \tag{8}$$

$$\vec{\phi}^{+}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \vec{a}^{+}_{\vec{k}}\, e^{-i(\vec{k}\vec{r} - kc_l t)} \tag{9}$$

with the condition

$$\int \vec{\phi}^{+}(\vec{r}, t)\, \vec{\phi}(\vec{r}, t)\, dV = n_0 + \sum_{\vec{k} \neq 0} \hat{a}^{+}_{\vec{k}}\, \hat{a}_{\vec{k}} = \hat{n}, \tag{10}$$

where $\vec{a}^{+}_{\vec{k}}$ and $\vec{a}_{\vec{k}}$ are, respectively, the Bose vector-operators of creation and annihilation for a free *sound particle* with the energy $\frac{\hbar^2 k^2}{2m}$, described by the vector $\vec{k}$ whose direction coincides with the direction $\vec{l}$ of a traveling longitudinal elastic wave; $\hat{n}$ is the operator of the total number of *sound particles*; $\hat{n}_0$ is the total number of *sound particles* at the condensate level with the wave vector $\vec{k} = 0$.

Thus, as is seen, the displacement vector $\vec{u}(\vec{r}, t)$ satisfies wave-equation (6) and in turn takes the form:

$$\vec{u}(\vec{r}, t) = \vec{u_0} + \frac{u_l}{\sqrt{V}} \sum_{\vec{k} \neq 0} \left(\vec{a}_{\vec{k}}\, e^{i(\vec{k}\vec{r} - kc_l t)} + \vec{a}^{+}_{\vec{k}}\, e^{-i(\vec{k}\vec{r} - kc_l t)}\right). \tag{11}$$

While investigating a superfluid liquid, Bogoliubov [8] separated the atoms of liquid helium ^{4}He in the condensate from those atoms filling the states above the condensate. In an analogous manner, we may consider the vector operator $\vec{a}_0 = \vec{l}\sqrt{n_0}$ and $\vec{a}^{+}_0 = \vec{l}\sqrt{n_0}$ as c-numbers (where $\vec{l}$ is the unit vector in the direction of propagation of the sound wave) within the approximation of a macroscopic number of *sound particles* in the condensate $n_0 \gg 1$. These assumptions lead to a broken Bose-symmetry law for *sound particles* in the condensate. To extend the concept of a broken Bose-symmetry law for *sound particles* in the condensate, we apply the definition of BEC of *sound particles* in the condensate as was postulated by the Penrose-Onsager for the definition of BEC of helium atoms [9]:

$$\lim_{n_0, n \to \infty} \frac{n_0}{n} = const. \tag{12}$$

On the other hand, we may observe that presence of *sound particles* filling the condensate level with the wave vector $\vec{k} = 0$ leads to the appearance of the constant displacement $\vec{u}_0 = \frac{2u_l \vec{l}\sqrt{n_0}}{\sqrt{V}}$ of the *sound particles*.

To find the normalization constant u_l, we introduce the following condition which allows us to suggest that at absolute zero all *sound particles* fill the condensate level $\vec{k} = 0$.

This reasoning implies that at $n_0 = n$ the constant displacement takes a maximal value $2d = \sqrt{|\vec{u}_0|^2}$ which represents the maximal distance between two neighboring *sound particles*. On the other hand, this distance is determined by the formula $d = \left(\frac{3V}{4\pi n}\right)^{\frac{1}{3}}$, which is in turn substituted into the expression $2d = \sqrt{|\vec{u}_0|^2}$. Then, consequently, we get to the normalization constant $u_l = 0.65\left(\frac{n}{V}\right)^{-\frac{5}{6}}$.

The condition of conservation of density at each point of the solid stipulates that

$$\rho_0 = \frac{MN}{V} = \frac{mn}{V}, \tag{13}$$

which represents a connection of the mass m and density ρ_0 of *sound particles* with the mass M and density ρ_0 of the liquid helium atoms with mass M.

Now, we consider the Hamiltonian operator $\hat{H}_l$ of a liquid [8]:

$$\hat{H}_l = \frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV + \frac{1}{2}\int\left(\frac{c_l\rho'}{\sqrt{\rho_0}}\right)^2 dV. \tag{14}$$

Substituting ρ' from (2) into (14), we obtain

$$\hat{H}_l = \frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV + \frac{\rho_0}{2}\int(c_l\,\mathrm{div}\;\vec{u})^2\,dV. \tag{15}$$

Using Dirac's approach in [10] for quantization of the electromagnetic field, we have:

$$\frac{\partial\vec{u}(\vec{r},t)}{\partial t} = -\frac{ic_l\vec{u}_l}{\sqrt{V}}\sum_{\vec{k}} k\left(\vec{a}_{\vec{k}}\,e^{-ikc_lt} - \vec{a}^{+}_{-\vec{k}}\,e^{ikc_lt}\right)e^{i\vec{k}\vec{r}} \tag{16}$$

as well as

$$\mathrm{div}\;\vec{u}(\vec{r},t) = \frac{i\vec{u}_l}{\sqrt{V}}\sum_{\vec{k}}\vec{k}\left(\vec{a}_{\vec{k}}\,e^{-ikc_lt} + \vec{a}^{+}_{-\vec{k}}\,e^{ikc_lt}\right)e^{i\vec{k}\vec{r}}. \tag{17}$$

Now, introducing (16) and (17) into (15) and using

$$\frac{1}{V}\int e^{i(\vec{k}_1+\vec{k}_2)\vec{r}} = \delta^3_{\vec{k}_1+\vec{k}_2},$$

we obtain the terms in the right side of the Hamiltonian of the system presented in (15):

$$\frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV = -\frac{\rho_0c_l^2u_l^2}{2}\sum_{\vec{k}}k^2\left(\vec{a}_{\vec{k}} - \vec{a}^{+}_{-\vec{k}}\right)\left(\vec{a}_{-\vec{k}} - \vec{a}^{+}_{\vec{k}}\right)$$

and

$$\frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV = \frac{\rho_0c_l^2u_l^2}{2}\sum_{\vec{k}}k^2\left(\vec{a}_{\vec{k}} + \vec{a}^{+}_{-\vec{k}}\right)\left(\vec{a}_{-\vec{k}} + \vec{a}^{+}_{\vec{k}}\right).$$

These expressions determine the reduced form of the Hamiltonian operator $\hat{H}_l$ by the form:

$$\hat{H}_l = 2\sum_{\vec{k}}\rho_0u_l^2c_l^2k^2\vec{a}^{+}_{\vec{k}}a_{\vec{k}}, \tag{18}$$

where u_l^2 is defined by the first term in the right side of (18) which represents the kinetic energy of a *sound particle* $\frac{\hbar^2k^2}{2m}$, if we suggest:

$$2\rho_0u_l^2c_l^2k^2 = \frac{\hbar^2k^2}{2m}. \tag{19}$$

Then,

$$u_l^2 = \frac{\hbar^2}{4c_l^2m\rho_0},$$

which allows one to determine the mass m of a *sound particle* using the value of the normalization constant $u_l = 0.65\left(\frac{n}{V}\right)^{-\frac{5}{6}}$ and (13):

$$m = \frac{\hbar}{c_l}\left(\frac{n}{V}\right)^{\frac{1}{3}}. \tag{20}$$

Thus, the Hamiltonian operator $\hat{H}_l$ describes an ideal Bose gas of a spinless *sound particles*:

$$\hat{H}_l = \sum_{\vec{k}}\frac{\hbar^2k^2}{2m}\vec{a}^{+}_{\vec{k}}a_{\vec{k}}. \tag{21}$$

3 Bose quasiparticles in solid

Now let us analyze quantization of a solid ^{4}He which consists of N atoms with the mass M confined in the volume V. Considering a solid ^{4}He as a continuum medium, we investigate the fluctuation motion of "solid particles" on the basis of hydrodynamics (where "solid particle" is determined as a very small volume V_0, in regard to the volume V of the solid ($V_0 \ll V$), which consists of a macroscopic number of ^{4}He atoms in solid).

To do the transition from quantum liquid to the solid ^{4}He, we introduce a concept as the fluctuation motion of "solid particles" or "solid points". In this respect, we remove such concept as a "lattice" of solid ^{4}He or such concept as an atoms, fixed in the knots of lattice because "solid particles" exist in any point of the solid. The motion of "solid particles" describe an elastic wave consisting of the *sound particles* with spin 1 which in turn are vibrated by the natural frequency Ω_l.

In this respect, we may express the vector displacement of a longitudinal ultrasonic wave $u_l(\vec{r},t)$ via the second quantization vector wave functions of one *sound particle* with spin 1. Then, Eqs. (8) and (9) take the forms:

$$\vec{\phi}(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{k},\sigma}\vec{a}_{\vec{k},\sigma}e^{i(\vec{k}\vec{r}-kc_lt)} \tag{22}$$

$$\vec{\phi}^{+}(\vec{r},t) = \frac{1}{\sqrt{V}}\sum_{\vec{k},\sigma}\vec{a}^{+}_{\vec{k},\sigma}e^{-i(\vec{k}\vec{r}-kc_lt)} \tag{23}$$

with condition

$$\int\phi^{+}(\vec{r},\sigma)\,\phi(\vec{r},\sigma)\,dV = n_0 + \sum_{\vec{k}\neq0,\sigma}\hat{a}^{+}_{\vec{k},\sigma}\hat{a}_{\vec{k},\sigma} = \hat{n}, \tag{24}$$

where a free sound particles have the mass m and the value of its spin z-component $\sigma = 0; \pm 1$. In this respect, the vector-operators $\vec{a}^{+}_{\vec{k},\sigma}$, $\vec{a}_{\vec{k},\sigma}$ satisfy the Bose commutation relations as:

$$\left[\hat{a}_{\vec{k},\sigma}, \hat{a}^{+}_{\vec{k}',\sigma'}\right] = \delta_{\vec{k},\vec{k}'} \cdot \delta_{\sigma,\sigma'}$$

$$[\hat{a}_{\vec{k},\sigma}, \hat{a}_{\vec{k}',\sigma'}] = 0$$

$$[\hat{a}^{+}_{\vec{k},\sigma}, \hat{a}^{+}_{\vec{k}',\sigma'}] = 0.$$

In this case, the Hamiltonian operator $\hat{H}$ of the solid ^{4}He is represented by the form:

$$\hat{H} = \frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV + \frac{1}{2}\int\left(\frac{c_l\rho'}{\sqrt{\rho_0}}\right)^2 dV + \\ + \frac{\rho_0}{2}\int(\Omega_l\vec{u}_l)^2\, dV, \tag{25}$$

where

$$\frac{\rho_0}{2}\int\left(\frac{\partial\vec{u}}{\partial t}\right)^2 dV = -\frac{\rho_0 c_l^2 u_l^2}{2}\sum_{\vec{k},\sigma} k^2\left(\vec{a}_{\vec{k},\sigma} - \vec{a}^{+}_{-\vec{k},\sigma}\right)\left(\vec{a}_{-\vec{k},\sigma} - \vec{a}^{+}_{\vec{k},\sigma}\right),$$

$$\frac{\rho_0}{2}\int(\operatorname{div}\vec{u})^2\, dV = \frac{\rho_0 c_l^2 u_l^2}{2}\sum_{\vec{k}} k^2\left(\vec{a}_{\vec{k},\sigma} + \vec{a}^{+}_{-\vec{k},\sigma}\right)\left(\vec{a}_{-\vec{k},\sigma} + \vec{a}^{+}_{\vec{k},\sigma}\right)$$

and

$$\frac{\rho_0}{2}\int(\Omega_l\vec{u}_l)^2\, dV = \frac{\rho_0\Omega^2 u_l^2}{2}\sum_{\vec{k}}\left(\vec{a}_{\vec{k},\sigma} + \vec{a}^{+}_{-\vec{k},\sigma}\right)\left(\vec{a}_{-\vec{k},\sigma} + \vec{a}^{+}_{\vec{k},\sigma}\right).$$

These expressions determine the reduced form of the Hamiltonian operator $\hat{H}$:

$$\hat{H}_l = \sum_{\vec{k}\neq 0,\sigma}\left(\frac{\hbar^2k^2}{2m} + mv^2\right)\vec{a}^{+}_{\vec{k},\sigma}a_{\vec{k},\sigma} + \\ + \frac{mv^2}{2}\sum_{\vec{k}\neq 0,\sigma}\left(\vec{a}^{+}_{-\vec{k},\sigma}\vec{a}^{+}_{\vec{k},\sigma} + \vec{a}_{\vec{k},\sigma}\vec{a}_{-\vec{k},\sigma}\right), \tag{26}$$

where we denote $v = \frac{\hbar\Omega_l}{\sqrt{2}mc_l}$, which in turn is the speed of sound particle in a solid.

For the evolution of the energy level, it is necessary to diagonalize the Hamiltonian $\hat{H}_l$, which can be accomplished by introducing the vector Bose-operators $\vec{b}^{+}_{\vec{k}}$ and $\vec{b}_{\vec{k}}$ [11]:

$$\vec{a}_{\vec{k},\sigma} = \frac{\vec{b}_{\vec{k},\sigma} + L_{\vec{k}}\,\vec{b}^{+}_{-\vec{k},\sigma}}{\sqrt{1 - L^2_{\vec{k}}}}, \tag{27}$$

where $L_{\vec{k}}$ is the unknown real symmetrical function of the wave vector $\vec{k}$.

By substituting (27) into (26), we obtain

$$\hat{H} = \sum_{\vec{k}\neq 0}\varepsilon_{\vec{k}}\vec{b}^{+}_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma}, \tag{28}$$

where $\vec{b}^{+}_{\vec{k},\sigma}$ and $\vec{b}_{\vec{k},\sigma}$ are the creation and annihilation operators of Bose quasiparticles with spin 1 with the energy:

$$\varepsilon_{\vec{k}} = \left[\left(\frac{\hbar^2k^2}{2m}\right)^2 + \hbar^2k^2v^2\right]^{1/2}. \tag{29}$$

In this context, the real symmetrical function $L_{\vec{k}}$ of the wave vector $\vec{k}$ is found to be

$$L^2_{\vec{k}} = \frac{\frac{\hbar^2k^2}{2m} + mv^2 - \varepsilon_{\vec{k}}}{\frac{\hbar^2k^2}{2m} + mv^2 + \varepsilon_{\vec{k}}}. \tag{30}$$

Thus, the average energy of the system takes the form:

$$\overline{\hat{H}} = \sum_{\vec{k}\neq 0}\varepsilon_{\vec{k}}\overline{\vec{b}^{+}_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma}}, \tag{31}$$

where $\overline{\vec{b}^{+}_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma}}$ is the average number of Bose quasiparticles with spin 1 with the wave vector $\vec{k}$ at the temperature T:

$$\overline{\vec{b}^{+}_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma}} = \frac{1}{e^{\frac{\varepsilon_{\vec{k}}}{kT}} - 1}. \tag{32}$$

Thus, we have found the spectrum of free quasiparticles with spin 1 which is similar to Bogoliubov's one [8]. In fact, the Hamiltonian of system (31) describes an ideal Bose gas consisting of phonons with spin 1 at a small wave number $k \ll \frac{2mv}{\hbar}$ but at $k \gg \frac{2mv}{\hbar}$ the Hamiltonian operator describes an ideal gas of *sound particles*.

4 BEC of *sound particles*

As opposed to London's postulation concerning BEC of atoms [12], we state that *sound particles* in the condensate define the superfluid component of solid ^{4}He. Consequently, statistical equilibrium equation (10) takes the following form:

$$n_{0,T} + \sum_{\vec{k}\neq 0}\overline{\vec{a}^{+}_{\vec{k},\sigma}\vec{a}_{\vec{k},\sigma}} = n, \tag{33}$$

where $\overline{\vec{a}^{+}_{\vec{k},\sigma}\vec{a}_{\vec{k},\sigma}}$ is the average number of *sound particles* with the wave vector $\vec{k}$ at the temperature T.

To find the form $\overline{\vec{a}^{+}_{\vec{k},\sigma}\vec{a}_{\vec{k},\sigma}}$, we use the linear transformation presented in (22):

$$\overline{\vec{a}^{+}_{\vec{k},\sigma}\vec{a}_{\vec{k},\sigma}} = \frac{1 + L^2_{\vec{p}}}{1 - L^2_{\vec{p}}}\overline{\vec{b}^{+}_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma}} + \\ + \frac{L_{\vec{k}}}{1 - L^2_{\vec{k}}}\left(\overline{\vec{b}^{+}_{\vec{k},\sigma}\vec{b}^{+}_{-\vec{k},\sigma}} + \overline{\vec{b}_{\vec{k},\sigma}\vec{b}_{-\vec{k},\sigma}}\right) + \frac{L^2_{\vec{k}}}{1 - L^2_{\vec{k}}}.$$

According to the Bloch-De-Dominicis theorem, we have

$$\overline{\vec{b}^{+}_{\vec{k},\sigma}\,\vec{b}^{+}_{-\vec{k},\sigma}} = \overline{\vec{b}_{\vec{k},\sigma}\vec{b}_{-\vec{k},\sigma}} = 0.$$

In this respect, the equation for the density of *sound particles* in the condensate takes the following form:

$$\frac{n_{0,T}}{V} = \frac{n}{V} - \frac{1}{V}\sum_{\vec{k}\neq 0,\sigma}\frac{L^2_{\vec{k}}}{1-L^2_{\vec{k}}} - \frac{1}{V}\sum_{\vec{k}\neq 0,\sigma}\frac{1+L^2_{\vec{k}}}{1-L^2_{\vec{k}}}\overline{\vec{b}^{+}_{\vec{k},\sigma}\vec{b}_{\vec{k},\sigma}}. \quad (34)$$

Obviously, at the lambda transition $T = T_\lambda$ the density of *sound particles* $\frac{n_{0,T_\lambda}}{V} = 0$. Hence, we note that the mass m and density $\frac{n}{V}$ of *sound particles* are expressed via the mass of ions M and density of ions $\frac{N}{V}$ when solving a system of two equations presented in (13) and (20):

$$\frac{n}{V} = \left(\frac{Mc_l}{\hbar}\frac{N}{V}\right)^{\frac{3}{4}} \quad (35)$$

and

$$m = \left(\frac{\hbar}{c_l}\right)^{\frac{3}{4}}\left(\frac{MN}{V}\right)^{\frac{1}{4}}. \quad (36)$$

At $T \to 0$ it follows $\overline{\vec{b}^{+}_{\vec{k},\sigma}\,\vec{b}_{\vec{k},\sigma}} = 0$. Then taking into account the coefficient with number 3 before integral on the right side of equation (34) because it reflects the value of spin z-component $\sigma = 0; \pm 1$, we obtain

$$\frac{n_{0,T}}{n} = 1 - \frac{m^3v^3}{\hbar^3\pi^2\frac{n}{V}}. \quad (37)$$

5 Conclusions

Thus, in this letter, we propose new model for solids which is different from the well-known models of Einstein and Debye because: 1) we suggest that the atoms are the Fermi particles which are absent in the Einstein and Debye models; 2) we remove such concept as lattice of solid by introducing a concept as the fluctuation motion of "solid particles" or "solid points". Thus, we deal with the "solid particle" which exist in any point of the solid; 3) In our model, we argue that the phonons in solid have spin 1 which is different from one presented by Einstein and Debye models; 4) in fact, in this letter, we first postulate that the superfluid component of a solid ^{4}He is determined by means of *sound particles* in the condensate as opposed to London's postulation concerning BEC of atoms [12]. Consequently, such reasoning allows us to consider the model of solid in a new light.

Submitted on April 18, 2011 / Accepted on April 25, 2011

References

1. Greywall D. S. Search in superfluidity in solid ^{4}He. *Physical Review B*, 1977, v. 16, 1291–1292.
2. Kim E., Chan M. H. W. Probable observation of a supersolid helium phase. *Nature*, 2004, v. 427, 225–227.
3. Diallo S. O. et al. Bose-Einstein Condensation in Solid ^{4}He.*Physical Review Letters*, 2007, v. 98, 205301.
4. Einstein A. Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Waerme. *Annalen der Physik*, 1907, v. 22, 180–190.
5. Debye P. Zur Theorie der spezifischen Waerme. *Annalen der Physik*, 1912, v. 39, 789–839.
6. Minasyan V. N., Samoilov V. N. Sound-Particles and Phonons with Spin 1. *Progress in Physics*, 2011, v. 1, 81–86; Minasyan V. N., Samoilov V. N. Charged Polaritons with Spin 1. *Progress in Physics*, 2011, v. 2, 7–12.
7. Landau L. D., Lifshiz E. M. Theory of Elasticity. *Theoretical Physics*, 1987, v. 11, 124–127.
8. Bogoliubov N. N. On the theory of superfludity. *Journal of Physics (USSR)*, 1974, v. 11, 23–32.
9. Penrose O., Onsager L. Bose-Einstein condensation and liquid helium. *Physical Review*, 1956, v. 104, 576–584.
10. Dirac P. A. M. The Principles of Quantum Mechanics. Clarendon press, Oxford, (1958).
11. Minasyan V. N., Samoilov V. N. Two Type Surface Polaritons Excited into Nanoholes in Metal Films. *Progress in Physics*, 2010, v. 2, 3–6.
12. London F. The λ-Phenomenon of Liquid Helium and the Bose-Einstein Degeneracy. *Nature*, 1938, v. 141, 643–644.

Fermion-Antifermion Asymmetry

Gunn Quznetsov

Chelyabinsk State University, Chelyabinsk, Ural, Russia. E-mail: gunn@mail.ru, quznets@yahoo.com

An event with positive energy transfers this energy photons which carries it on recorders observers. Observers know that this event occurs, not before it happens. But events with negative energy should absorb this energy from observers. Consequently, observers know that this event happens before it happens. Since time is irreversible then only the events with positive energy can occur. In single-particle states, events with a fermion have positive energy and occurrences with an antifermion have negative energy. In double-particle states, events with pair of antifermions have negative energy and events with pair of fermions and with fermion-antifermion pair have positive energy.

1 Introduction

Let t, x_1, x_2,x_3 be real numbers, and let $\mathbf{x} := \langle x_1, x_2, x_3\rangle$.

Let $\mathcal{A}$ be some pointlike event.

Let $\varphi(t,\mathbf{x})$ be a 4×1-complex matrix such that

$$\varphi^{\dagger}(t,\mathbf{x})\varphi(t,\mathbf{x}) = \rho(t,\mathbf{x}) \tag{1}$$

where $\rho(t,\mathbf{x})$ is the probability density of $\mathcal{A}$.

Let* $\rho(t,\mathbf{x}) = 0$ if $t > \frac{\pi \mathrm{c}}{\mathrm{h}}$ and/or $|\mathbf{x}| > \frac{\pi \mathrm{c}}{\mathrm{h}}$.

In that case $\varphi(t,\mathbf{x})$ obeys some generalization of the Dirac equation [1]. The Dirac equation for free fermion does have the following form:

$$\left(\frac{1}{\mathrm{c}}\frac{\partial}{\partial t} - \sum_{s=1}^{3}\beta^{[s]}\frac{\partial}{\partial x_s} - \mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}n\gamma^{[0]}\right)\varphi(t,\mathbf{x}) = 0.$$

Here n is a natural number and

$$\beta^{[1]} := \begin{bmatrix} 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & -1\\ 0 & 0 & -1 & 0 \end{bmatrix}, \beta^{[2]} := \begin{bmatrix} 0 & -\mathrm{i} & 0 & 0\\ \mathrm{i} & 0 & 0 & 0\\ 0 & 0 & 0 & \mathrm{i}\\ 0 & 0 & -\mathrm{i} & 0 \end{bmatrix},$$

$$\beta^{[3]} := \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}, \gamma^{[0]} := \begin{bmatrix} 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

In this case operator $\widehat{H_0}$ is *the free Dirac Hamiltonian* if

$$\widehat{H_0} := \mathrm{c}\left(\sum_{s=1}^{3}\beta^{[s]}\mathrm{i}\frac{\partial}{\partial x_s} + \frac{\mathrm{h}}{\mathrm{c}}n\gamma^{[0]}\right).$$

Let $\mathbf{k}$ be a vector $\langle k_1, k_2, k_3\rangle$ where k_s are integer numbers and let

$$\omega(\mathbf{k}) := \sqrt{k_1^2 + k_2^2 + k_3^2 + n^2}$$

where n is a natural number.

*$\mathrm{c} := 299792458, \mathrm{h} := 6.6260755^{-34}$

Let

$$e_1(\mathbf{k}) := \frac{1}{2\sqrt{\omega(\mathbf{k})(\omega(\mathbf{k})+n)}}\begin{bmatrix} \omega(\mathbf{k})+n+k_3\\ k_1+\mathrm{i}k_2\\ \omega(\mathbf{k})+n-k_3\\ -k_1-\mathrm{i}k_2 \end{bmatrix},$$

$$e_2(\mathbf{k}) := \frac{1}{2\sqrt{\omega(\mathbf{k})(\omega(\mathbf{k})+n)}}\begin{bmatrix} k_1-\mathrm{i}k_2\\ \omega(\mathbf{k})+n-k_3\\ -k_1-\mathrm{i}k_2\\ \omega(\mathbf{k})+n+k_3 \end{bmatrix},$$

$$e_3(\mathbf{k}) := \frac{1}{2\sqrt{\omega(\mathbf{k})(\omega(\mathbf{k})+n)}}\begin{bmatrix} -\omega(\mathbf{k})-n+k_3\\ k_1+\mathrm{i}k_2\\ \omega(\mathbf{k})+n+k_3\\ k_1+\mathrm{i}k_2 \end{bmatrix},$$

$$e_4(\mathbf{k}) := \frac{1}{2\sqrt{\omega(\mathbf{k})(\omega(\mathbf{k})+n)}}\begin{bmatrix} k_1-\mathrm{i}k_2\\ -\omega(\mathbf{k})-n-k_3\\ k_1-\mathrm{i}k_2\\ \omega(\mathbf{k})+n-k_3 \end{bmatrix}.$$

In that case, functions
$e_1(\mathbf{k})(2\mathrm{c}/\mathrm{h})^{3/2}\exp(-\mathrm{i}(\mathrm{h}/\mathrm{c})\mathbf{kx})$ and
$e_2(\mathbf{k})(2\mathrm{c}/\mathrm{h})^{3/2}\exp(-\mathrm{i}(\mathrm{h}/\mathrm{c})\mathbf{kx})$
are eigenvectors of $\widehat{H_0}$ with eigenvalues $(+\mathrm{h}\omega(\mathbf{k}))$,
and functions
$e_3(\mathbf{k})(2\mathrm{c}/\mathrm{h})^{3/2}\exp(-\mathrm{i}(\mathrm{h}/\mathrm{c})\mathbf{kx})$ and
$e_4(\mathbf{k})(2\mathrm{c}/\mathrm{h})^{3/2}\exp(-\mathrm{i}(\mathrm{h}/\mathrm{c})\mathbf{kx})$
are eigenvectors of $\widehat{H_0}$ with eigenvalues $(-\mathrm{h}\omega(\mathbf{k}))$.

2 Single-Particle States

Let $\mathfrak{H}$ be some unitary space. Let $\widetilde{0}$ be the zero element of $\mathfrak{H}$. That is any element $\widetilde{F}$ of $\mathfrak{H}$ obeys to the following conditions:

$$0\widetilde{F} = \widetilde{0}, \widetilde{0} + \widetilde{F} = \widetilde{F}, \widetilde{0}^{\dagger}\widetilde{F} = \widetilde{F}, \widetilde{0}^{\dagger} = \widetilde{0}.$$

Let $\widehat{0}$ be the zero operator on $\mathfrak{H}$. That is any element $\widetilde{F}$ of

$\mathfrak{H}$ obeys to the following condition:

$\widetilde{0F} = 0\widetilde{F}$, and if $\widehat{b}$ is any operator on $\mathfrak{H}$ then

$\widehat{0} + \widehat{b} = \widehat{b} + \widehat{0} = \widehat{b}, \widetilde{0b} = \widetilde{b0} = \widehat{0}.$

Let $\widehat{1}$ be the identy operator on $\mathfrak{H}$. That is any element $\widetilde{F}$ of $\mathfrak{H}$ obeys to the following condition:
$\widetilde{1F} = 1\widetilde{F} = \widetilde{F}$, and if $\widehat{b}$ is any operator on $\mathfrak{H}$ then $\widehat{1b} = \widehat{b1} = \widehat{b}$.

Let linear operators $b_{s,\mathbf{k}}$ ($s \in \{1,2,3,4\}$) act on all elements of this space. And let these operators fulfill the following conditions:

$$\left\{b^{\dagger}_{s,\mathbf{k}}, b_{s',\mathbf{k}'}\right\} := b^{\dagger}_{s,\mathbf{k}} b_{s',\mathbf{k}'} + b_{s',\mathbf{k}'} b^{\dagger}_{s,\mathbf{k}} = \left(\frac{\mathrm{h}}{2\pi\mathrm{c}}\right)^3 \delta_{\mathbf{k},\mathbf{k}'}\delta_{s,s'}\widehat{1},$$

$$\{b_{s,\mathbf{k}}, b_{s',\mathbf{k}'}\} = b_{s,\mathbf{k}} b_{s',\mathbf{k}'} + b_{s',\mathbf{k}'} b_{s,\mathbf{k}} = \left\{b^{\dagger}_{s,\mathbf{k}}, b^{\dagger}_{s',\mathbf{k}'}\right\} = \widehat{0}.$$

Hence,

$$b_{s,\mathbf{k}} b_{s,\mathbf{k}} = b^{\dagger}_{s,\mathbf{k}} b^{\dagger}_{s,\mathbf{k}} = \widehat{0}.$$

There exists element $\widetilde{F}_0$ of $\mathfrak{H}$ such that $\widetilde{F}^{\dagger}_0 \widetilde{F}_0 = 1$ and for any $b_{s,\mathbf{k}}$: $b_{s,\mathbf{k}}\widetilde{F}_0 = \widetilde{0}$. Hence, $\widetilde{F}^{\dagger}_0 b^{\dagger}_{s,\mathbf{k}} = \widetilde{0}$.

Let

$$\psi_s(\mathbf{x}) := \sum_{\mathbf{k}} \sum_{r=1}^{4} b_{r,\mathbf{k}} e_{r,s}(\mathbf{k}) \exp\left(-\mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}\mathbf{k}\mathbf{x}\right).$$

Because

$$\sum_{r=1}^{4} e_{r,s}(\mathbf{k}) e_{r,s'}(\mathbf{k}) = \delta_{s,s'}$$

and

$$\sum_{\mathbf{k}} \exp\left(-\mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}\mathbf{k}(\mathbf{x}-\mathbf{x}')\right) = \left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^3 \delta(\mathbf{x}-\mathbf{x}')$$

then

$$\begin{aligned}\left\{\psi^{\dagger}_s(\mathbf{x}), \psi_{s'}(\mathbf{x}')\right\} &:= \psi^{\dagger}_s(\mathbf{x})\psi_{s'}(\mathbf{x}') + \psi_{s'}(\mathbf{x}')\psi^{\dagger}_s(\mathbf{x}) \\ &= \delta(\mathbf{x}-\mathbf{x}')\delta_{s,s'}\widehat{1}.\end{aligned}$$

And these operators obey the following conditions:

$$\psi_s(\mathbf{x})\widetilde{F}_0 = \widetilde{0}, \{\psi_s(\mathbf{x}), \psi_{s'}(\mathbf{x}')\} = \left\{\psi^{\dagger}_s(\mathbf{x}), \psi^{\dagger}_{s'}(\mathbf{x}')\right\} = \widehat{0}.$$

Hence,

$$\psi_s(\mathbf{x})\psi_{s'}(\mathbf{x}') = \psi^{\dagger}_s(\mathbf{x})\psi^{\dagger}_{s'}(\mathbf{x}') = \widehat{0}.$$

Let

$$\Psi(t,\mathbf{x}) := \sum_{s=1}^{4} \varphi_s(t,\mathbf{x})\psi^{\dagger}_s(\mathbf{x})\widetilde{F}_0.$$

These functions obey the following condition:

$$\Psi^{\dagger}(t,\mathbf{x}')\Psi(t,\mathbf{x}) = \varphi^{\dagger}(t,\mathbf{x}')\varphi(t,\mathbf{x})\delta(\mathbf{x}-\mathbf{x}').$$

Hence,

$$\int d\mathbf{x}' \cdot \Psi^{\dagger}(t,\mathbf{x}')\Psi(t,\mathbf{x}) = \rho(t,\mathbf{x}).$$

Let a Fourier series of $\varphi_s(t,\mathbf{x})$ has the following form:

$$\varphi_s(t,\mathbf{x}) = \sum_{\mathbf{p}} \sum_{r=1}^{4} c_r(t,\mathbf{p}) e_{r,s}(\mathbf{p}) \exp\left(-\mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}\mathbf{p}\mathbf{x}\right).$$

In that case:

$$\underline{\Psi}(t,\mathbf{p}) := \left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^3 \sum_{r=1}^{4} c_r(t,\mathbf{p}) b^{\dagger}_{r,\mathbf{p}}\widetilde{F}_0.$$

If

$$\mathcal{H}_0(\mathbf{x}) := \psi^{\dagger}(\mathbf{x})\widehat{H}_0\psi(\mathbf{x}) \tag{2}$$

then $\mathcal{H}_0(\mathbf{x})$ is called a Hamiltonian $\widehat{H}_0$ density.

Because

$$\widehat{H}_0\varphi(t,\mathbf{x}) = \mathrm{i}\frac{\partial}{\partial t}\varphi(t,\mathbf{x})$$

then

$$\int d\mathbf{x}' \cdot \mathcal{H}_0(\mathbf{x}')\Psi(t,\mathbf{x}) = \mathrm{i}\frac{\partial}{\partial t}\Psi(t,\mathbf{x}). \tag{3}$$

Therefore, if

$$\widehat{\mathbb{H}} := \int d\mathbf{x}' \cdot \mathcal{H}_0(\mathbf{x}')$$

then $\widehat{\mathbb{H}}$ acts similar to the Hamiltonian on space $\mathfrak{H}$.
And if

$$E_{\Psi}\left(\widetilde{F}_0\right) := \sum_{\mathbf{p}} \underline{\Psi}^{\dagger}(t,\mathbf{p})\widehat{\mathbb{H}}\underline{\Psi}(t,\mathbf{p})$$

then $E_{\Psi}\left(\widetilde{F}_0\right)$ is an energy of Ψ on vacuum $\widetilde{F}_0$.

Operator $\widehat{\mathbb{H}}$ obeys the following condition:

$$\widehat{\mathbb{H}} = \left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^3 \sum_{\mathbf{k}} \mathrm{h}\omega(\mathbf{k}) \left(\sum_{r=1}^{2} b^{\dagger}_{r,\mathbf{k}} b_{r,\mathbf{k}} - \sum_{r=3}^{4} b^{\dagger}_{r,\mathbf{k}} b_{r,\mathbf{k}}\right).$$

This operator is not positive defined and in this case

$$E_{\Psi}\left(\widetilde{F}_0\right) = \left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^3 \sum_{\mathbf{p}} \mathrm{h}\omega(\mathbf{p}) \left(\sum_{r=1}^{2} |c_r(t,\mathbf{p})|^2 - \sum_{r=3}^{4} |c_r(t,\mathbf{p})|^2\right).$$

This problem is usually solved in the following way [2, p. 54]:

Let:

$$\begin{aligned} v_1(\mathbf{k}) &:= \gamma^{[0]} e_3(\mathbf{k}), \\ v_2(\mathbf{k}) &:= \gamma^{[0]} e_4(\mathbf{k}), \\ d_{1,\mathbf{k}} &:= -b^{\dagger}_{3,-\mathbf{k}}, \\ d_{2,\mathbf{k}} &:= -b^{\dagger}_{4,-\mathbf{k}}. \end{aligned}$$

In that case:

$$\begin{aligned} e_3(\mathbf{k}) &= -v_1(-\mathbf{k}), \\ e_4(\mathbf{k}) &= -v_2(-\mathbf{k}), \\ b_{3,\mathbf{k}} &= -d^{\dagger}_{1,-\mathbf{k}}, \\ b_{4,\mathbf{k}} &= -d^{\dagger}_{2,-\mathbf{k}}. \end{aligned}$$

Therefore,

$$\begin{aligned} \psi_s(\mathbf{x}) \;:\; &= \sum_{\mathbf{k}} \sum_{r=1}^{2} \Big(b_{r,\mathbf{k}} e_{r,s}(\mathbf{k}) \exp\left(-\mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}\mathbf{kx}\right) + \\ &+ d^{\dagger}_{r,\mathbf{k}} v_{r,s}(\mathbf{k}) \exp\left(\mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}\mathbf{kx}\right)\Big) \end{aligned}$$

$$\begin{aligned} \widehat{\mathbb{H}} &= \left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^3 \sum_{\mathbf{k}} \mathrm{h}\omega(\mathbf{k}) \sum_{r=1}^{2} \left(b^{\dagger}_{r,\mathbf{k}} b_{r,\mathbf{k}} + d^{\dagger}_{r,\mathbf{k}} d_{r,\mathbf{k}} \right) \\ &\quad -2 \sum_{\mathbf{k}} \mathrm{h}\omega(\mathbf{k}) \widehat{1}. \end{aligned}$$

The first term on the right side of this equality is positive defined. This term is taken as the desired Hamiltonian. The second term of this equality is infinity constant. And this infinity is deleted (?!) [2, p. 58]

But in this case $d_{r,\mathbf{k}}\widetilde{F}_0 \neq \widetilde{0}$. In order to satisfy such condition, the vacuum element $\widetilde{F}_0$ must be replaced by the following:

$$\widetilde{F}_0 \to \widetilde{\Phi}_0 := \prod_{\mathbf{k}} \prod_{r=3}^{4} \left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^3 b^{\dagger}_{r,\mathbf{k}} \widetilde{F}_0.$$

But in this case:

$$\psi_s(\mathbf{x})\,\widetilde{\Phi}_0 \neq \widetilde{0}.$$

And condition (3) isn't carried out.

In order to satisfy such condition, operators $\psi_s(\mathbf{x})$ must be replaced by the following:

$$\begin{aligned} &\psi_s(\mathbf{x}) \to \phi_s(\mathbf{x}) := \\ &:= \sum_{\mathbf{k}} \sum_{r=1}^{2} \Big(b_{r,\mathbf{k}} e_{r,s}(\mathbf{k}) \exp\left(-\mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}\mathbf{kx}\right) + d_{r,\mathbf{k}} v_r(\mathbf{k}) \exp\left(\mathrm{i}\frac{\mathrm{h}}{\mathrm{c}}\mathbf{kx}\right)\Big). \end{aligned}$$

Hence,

$$\begin{aligned} \widehat{\mathbb{H}} &= \int d\mathbf{x} \cdot \mathcal{H}(\mathbf{x}) = \int d\mathbf{x} \cdot \phi^{\dagger}(\mathbf{x}) \widehat{H}_0 \phi(\mathbf{x}) = \\ &= \left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^3 \sum_{\mathbf{k}} \mathrm{h}\omega(\mathbf{k}) \sum_{r=1}^{2} \left(b^{\dagger}_{r,\mathbf{k}} b_{r,\mathbf{k}} - d^{\dagger}_{r,\mathbf{k}} d_{r,\mathbf{k}} \right). \end{aligned}$$

And again we get negative energy.

Let's consider the meaning of such energy: An event with positive energy transfers this energy photons which carries it on recorders observers. Observers know that this event occurs, not before it happens. But event with negative energy should absorb this energy from observers. Consequently, observers know that this event happens before it happens. This contradicts Theorem 3.4.2 [3]. Therefore, events with negative energy do not occur.

Hence, over vacuum $\widetilde{\Phi}_0$ single fermions can exist, but there are no single antifermions.

3 Two-Particle States

A two-particle state is defined the following field operator [4]:

$$\psi_{s_1,s_2}(\mathbf{x},\mathbf{y}) := \begin{vmatrix} \phi_{s_1}(\mathbf{x}) & \phi_{s_2}(\mathbf{x}) \\ \phi_{s_1}(\mathbf{y}) & \phi_{s_2}(\mathbf{y}) \end{vmatrix}.$$

In that case:

$$\widehat{\mathbb{H}} = 2\mathrm{h}\left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^6 \left(\widehat{\mathbb{H}}_a + \widehat{\mathbb{H}}_b\right)$$

where

$$\begin{aligned} \widehat{\mathbb{H}}_a \;:\; &= \sum_{\mathbf{k}} \sum_{\mathbf{p}} (\omega(\mathbf{k}) - \omega(\mathbf{p})) \sum_{r=1}^{2} \sum_{j=1}^{2} \times \\ &\times \Big\{ v^{\dagger}_j(-\mathbf{k}) v_j(-\mathbf{p}) e^{\dagger}_r(\mathbf{p}) e_r(\mathbf{k}) \times \\ &\times \left(+ b^{\dagger}_{r,\mathbf{p}} d^{\dagger}_{j,-\mathbf{k}} d_{j,-\mathbf{p}} b_{r,\mathbf{k}} \right) + \\ &+ \left(+ d^{\dagger}_{r,-\mathbf{p}} b^{\dagger}_{j,\mathbf{k}} b_{j,\mathbf{k}} d_{r,-\mathbf{p}} \right) + \\ &+ v^{\dagger}_j(-\mathbf{p}) v_j(-\mathbf{k}) e^{\dagger}_r(\mathbf{k}) e_r(\mathbf{p}) \times \\ &\times \left(- b^{\dagger}_{r,\mathbf{k}} d^{\dagger}_{j,-\mathbf{p}} d_{j,-\mathbf{k}} b_{r,\mathbf{p}} \right) + \\ &+ \left(- b^{\dagger}_{r,\mathbf{p}} d^{\dagger}_{j,-\mathbf{k}} d_{j,-\mathbf{k}} b_{r,\mathbf{p}} \right) \Big\} \end{aligned}$$

and

$$\begin{aligned} \widehat{\mathbb{H}}_b \;:\; &= \sum_{\mathbf{k}} \sum_{\mathbf{p}} (\omega(\mathbf{k}) + \omega(\mathbf{p})) \sum_{r=1}^{2} \sum_{j=1}^{2} \times \\ &\times \Big\{ v^{\dagger}_j(-\mathbf{p}) v_j(-\mathbf{k}) v^{\dagger}_r(-\mathbf{k}) v_r(-\mathbf{p}) \times \\ &\times \left(- d^{\dagger}_{r,-\mathbf{k}} d^{\dagger}_{j,-\mathbf{p}} d_{j,-\mathbf{k}} d_{r,-\mathbf{p}} \right) + \\ &+ \left(- d^{\dagger}_{r,-\mathbf{p}} d^{\dagger}_{j,-\mathbf{k}} d_{j,-\mathbf{k}} d_{r,-\mathbf{p}} \right) \\ &+ e^{\dagger}_r(\mathbf{k}) e_r(\mathbf{p}) e^{\dagger}_j(\mathbf{p}) e_j(\mathbf{k}) \times \\ &\times \left(+ b^{\dagger}_{r,\mathbf{k}} b^{\dagger}_{j,\mathbf{p}} b_{j,\mathbf{k}} b_{r,\mathbf{p}} \right) + \\ &+ \left(+ b^{\dagger}_{r,\mathbf{p}} b^{\dagger}_{j,\mathbf{k}} b_{j,\mathbf{k}} b_{r,\mathbf{p}} \right) \Big\}. \end{aligned}$$

If velocities are small then the following formula is fair.

$$\widehat{\mathbb{H}} = 4\mathrm{h}\left(\frac{2\pi\mathrm{c}}{\mathrm{h}}\right)^6 \left(\widehat{\mathbb{H}}_a + \widehat{\mathbb{H}}_b\right)$$

where

$$\widehat{\mathbb{H}}_a \;:=\; \sum_{\mathbf{k}}\sum_{\mathbf{p}} \left(\omega\left(\mathbf{k}\right)-\omega\left(\mathbf{p}\right)\right)\times \\ \times\sum_{r=1}^{2}\sum_{j=1}^{2}\left(d^{\dagger}_{j,-\mathbf{p}}b^{\dagger}_{r,\mathbf{k}}b_{r,\mathbf{k}}d_{j,-\mathbf{p}}-b^{\dagger}_{j,\mathbf{p}}d^{\dagger}_{r,-\mathbf{k}}d_{r,-\mathbf{k}}b_{j,\mathbf{p}}\right)$$

and

$$\widehat{\mathbb{H}}_b \;:=\; \sum_{\mathbf{k}}\sum_{\mathbf{p}} \left(\omega\left(\mathbf{k}\right)+\omega\left(\mathbf{p}\right)\right)\times \\ \times\sum_{j=1}^{2}\sum_{r=1}^{2}\left(b^{\dagger}_{j,\mathbf{p}}b^{\dagger}_{r,\mathbf{k}}b_{r,\mathbf{k}}b_{j,\mathbf{p}}-d^{\dagger}_{j,-\mathbf{p}}d^{\dagger}_{r,-\mathbf{k}}d_{r,-\mathbf{k}}d_{j,-\mathbf{p}}\right).$$

Therefore, in any case events with pairs of fermions and events with fermion-antifermion pairs can occur, but events with pairs of antifermions can not happen.

4 Conclusion

Therefore, an antifermion can exist only with a fermion.

Submitted on April 27, 2011 / Accepted on April 30, 2011

References

1. Quznetsov G. 4×1-Marix Functions and Dirac's Equation, *Progress in Physics*, 2009, v. 2, 96–106.
2. For instance, Peskin M. E., Schroeder D. V. An Introduction to Quantum Field Theory, Perseus Books Publishing, L.L.C., 1995.
3. Quznetsov G. Logical Foundation of Theoretical Physics, Nova Sci. Publ., 2006, 34–35.
4. For instance, Ziman J. M. Elements of Advanced Quantum Theory, Cambridge, 1969, 32.

An Insight into Planck's Units: Explaining the Experimental–Observations of Lack of Quantum Structure of Space-Time

Hasmukh K. Tank

Indian Space Research Organization 22/693 Krishna Dham-2, Vejalpur, Ahmedabad-380015 India.
E-mail: tank.hasmukh@rediffmail.com

This letter presents an insight into Planck's natural-units, that they are geometric-mean-values of astronomical-quantities, like total-mass of the universe M_0 and mass corresponding to Hubble's-constant $\left(hH_0/c^2\right)$, providing a theoretical support to the observational findings of Ragazzoni, R., Turatto, M. & Gaessler [Astrophysical Journal,587, L1–L4], Lieu, R. & Hillman, L.W [Astrophysical Journal, 585, L77–L80] and a news item published in Nature [Published on line on 31 March 2003 Nature DOI 10.1038/news030324-13] that there is no observational evidence for the quantum structure of space-time. Physicists have been expecting unification of gravitational and electric forces at Planck's energy; so they wanted to experimentally create a pair of particles whose gravitational-radius is equal to their Compton-wavelength. Whereas this paper shows that in nature there exists a "pair of unequal masses" which satisfies the condition of equality of gravitational and electrostatic potential-energies of the pair. If the universe with its total-mass M_0 and a particle of mass hH_0/c^2 both are electrically charged bodies, then the strengths of electric force and gravitational-force experienced by them will be equal. It is also pointed-out here that P.A.M. Dirac's observation of recurrences of the large-number 10^{40} and its explanation proposed by Tank [Proceedings of Indian National Sci. Acad. A, Vol. 63, No. 6, 469–474 (1997)] in 1997, by Sidharth [arXiv:gen-ph/0509026] in 2005, and by Funkhouser [arXiv:gen-ph/0611115] in 2006, should be viewed as attempts in search of natural system of units; and the recurrences $R_0/r_e = e^2/Gm_e$, $m_p = \left[M_0/m_p\right]^{1/2}$ should be taken more seriously than a mere coincidence, because its explanation by Tank also helped explaining the recurrences of the critical-acceleration of MOND noticed by Sivaram [Astrophys. and Space Sci. 215, (1994), 185–189].

1 Introduction

It has been realized by physicists since long that the conventional system of units, like meter, kilogram and second are arbitrarily chosen units; they do not correspond with any fundamental physical quantities; so we find it difficult to observe any regular pattern. Max Plank proposed a set of natural-units. Physicists have been expecting unification of gravitational and electric forces at the energies where protons attain the masses close to Planck's-mass. Large Hadrons Collider [LHC] was expected to yield some interesting results, because protons were to attain Planck's mass. It was believed that space and time are quantized; Planck-length is the "least-count" for "space" and Planck's unit of "time" is the "least-count" for "time". Whereas this letter shows that Planck's units are statistical-quantities, they are geometric-mean-values of the astronomical-quantities like total-mass of the universe M_0 and mass corresponding to Hubble's constant (hH_0/c^2).

(i) Planck's length L^* is a geometric-mean of: Gravitational-Radius corresponding to total mass of the universe M_0 and Compton-wavelength corresponding to the total-mass M_0 of the universe, i.e.

$$L^* = \left[\left(GM_0/c^2\right)(h/M_0c)\right]^{1/2}.$$

Also, L^* is a geometric-mean of: gravitational-radius of the universe and that of the lightest-particle of mass $\left(hH_0/c^2\right)$. L^* is also a geometric-mean of Compton-wavelengths of M_0 and $\left(hH_0/c^2\right)$.

(ii) Planck's unit of time T^* is a geometric-mean of age-of-the-universe T_0 and the period corresponding the total mass of the universe h/M_0c^2.

(iii) Planck's unit of mass M^* is a geometric-mean of total-mass-of the-universe M_0 and mass-of-the-lightest-particle. So, this letter provides a theoretical explanation for the experimental observations by Ragazzoni et al [1] and Lieu et. al. [2] that there is no evidence for quantum structure of space-time.

(iv) The total mass of the universe M_0 and mass corresponding to Hubble's constant $\left(hH_0/c^2\right)$ form an interesting pair, that: Gravitational-Radius corresponding to total-mass of the universe is equal to Compton-wavelength of the lightest particle, of mass hH_0/c^2.

(v) Gravitational-radius of the lightest particle is equal to the Compton-wavelength of the total-mass of the universe, M_0. Physicists have been trying to generate a pair of particles of equal masses whose gravitational-radius is equal to their Compton-wavelength. But in nature, there exists a pair of unequal masses which satisfies the condition for unification of forces, that their gravitational-potential-energy should be equal to the electrostatic-potential-energy. So this pair is expected to provide some clue to a deeper understanding needed for unification of gravitational and electric forces.

It is also pointed-out here that P. A. M. Dirac's observation of recurrences of the large-number 10^{40} and its explanation proposed by Tank [4] in 1997, by Sidharth [5] in 2005, and by Funkhouser [6] in 2006, should be viewed as attempts in search of natural system of units; and the recurrences $R_0/r_e = e^2/Gm_e$, $m_p = \left[M_0/m_p\right]^{1/2}$ should be taken more seriously than a mere coincidence, because their explanation by Tank also helped explaining the recurrences of the critical-acceleration of MOND noticed by Sivaram [7] and led to further conclusions discussed in the references [8–10].

2 The Derivations

(i) Gravitational-Radius of the universe is equal to Compton-wavelength of the lightest particle, of mass hH_0/c^2:

The gravitational-radius-of-the-universe $R_0 = GM_0/c^2$; Here M_0 is total-mass of the universe. And Compton-wavelength of the lightest-particle of mass hH_0/c^2 ; where H_0 is Hubble's constant, is:

$$\begin{aligned} & h/\left(hH_0/c^2\right)c \\ \text{i.e.} \quad &= c/H_0, \\ \text{i.e.} \quad &= R_0, \\ \text{i.e.} \quad &= GM_0/c^2. \end{aligned}$$

(ii) Gravitational-radius of the lightest particle is equal to Compton-wavelength of the total-mass of the universe, M_0.

$$\begin{aligned} \text{i.e.} \quad &= G\left(hH_0/c^2\right)/c^2, \\ \text{i.e} \quad &= GhH_0/c^4, \\ \text{i.e.} \quad &= GhH_0/GH_0M_0c \end{aligned}$$

(Because $GH_0M_0 = c^3$, based on this author's previous work [4]), i.e.= h/M_0c which is the Compton-wavelength corresponding to the total-mass-of-the-universe.

(iii-a) Planck's length L^* is a geometric-mean of: Gravitational-Radius of the universe and Compton-wavelength corresponding to the total-mass of the universe:

$$\begin{aligned} \text{i.e. } L^* &= \left[\left(GM_0/c^2\right)(h/M_0c)\right]^{1/2}, \\ \text{i.e.} \quad &= \left[hG/c^3\right]^{1/2}. \end{aligned}$$

Similarly, Planck's length is a geometric-mean of gravitational-radius and Compton-wavelengths of every particle of any mass.

(iii-b) Planck's length L^* is also a geometric-mean of: gravitational-radius of the universe and that of the lightest-particle of mass hH_0/c^2:

That is:

$$\begin{aligned} & \left[\left(GM_0/c^2\right)\left(GhH_0/c^4\right)\right]^{1/2}, \\ \text{i.e.} \quad &= \left[G^2M_0hH_0/GM_0H_0c^3\right]^{1/2} \end{aligned}$$

(Because $GH_0M_0 = c^3$, based on this author's previous work [4]),

$$\begin{aligned} \text{i.e.} \quad &= \left[hG/c^3\right]^{1/2}, \\ \text{i.e.} \quad &= L^*. \end{aligned}$$

(iii-c) L^* is also a geometric-mean of Compton-wavelengths of M_0 and $\left(hH_0/c^2\right)$:

That is:

$$\begin{aligned} & \left[(h/M_0c)\left\{h/\left(hH_0/c^2\right)c\right\}\right]^{1/2}, \\ \text{i.e.} \quad &= \left[(h/M_0c)\,(c/H_0)\right]^{1/2}, \\ \text{i.e.} \quad &= \left[(h/M_0c)\,(R_0)\right]^{1/2}, \\ \text{i.e.} \quad &= \left[(h/M_0c)\left(GM_0/c^2\right)\right]^{1/2}, \\ \text{i.e.} \quad &= \left[hG/c^3\right]^{1/2}, \\ \text{i.e.} \quad &= L^*. \end{aligned}$$

The references [1–3] also lead to a conclusion that nothing very special is observed at Planck length; there is no evidence for any quantum structure of space-time. This paper has shown that Planck-length is a statistical-quantity, a geometric-mean-value, not a length of any fundamental-entity.

(iv) Planck's unit of time T^* is a geometric-mean of age-of-the-universe and the period corresponding the total-mass of the universe h/M_0c^2

Age-of-the-universe $T_0 = 1/H_0$.

So the product of the two periods is:

$$\begin{aligned} & (1/H_0)\left(h/M_0c^2\right), \\ \text{i.e.} \quad &= h/H_0M_0c^2, \\ \text{i.e.} \quad &= hG/c^5 \end{aligned}$$

(Because $GH_0M_0 = c^3$, based on this author's previous work [4])

$$\begin{aligned} \text{i.e.} \quad &= T^{*2}, \\ \text{i.e. } T^* &= \left[(T_0)\left(h/M_0c^2\right)\right]^{1/2}. \end{aligned}$$

(v) Planck's unit of mass M^* is a geometric-mean of total-mass-of the-universe M_0 and mass-of-the-lightest--particle :

$$\begin{aligned} \text{i.e.} \quad &= \left[(M_0)\left(hH_0/c^2\right)\right]^{1/2}, \\ \text{i.e.} \quad &= \left[M_0hH_0c/c^3\right]^{1/2}, \\ \text{i.e.} \quad &= \left[M_0hH_0c/GM_0H_0\right]^{1/2} \end{aligned}$$

(Because $GH_0M_0 = c^3$, based on this author's previous work [4]),

$$\text{i.e.} \quad = \quad [hc/G]^{1/2},$$
$$\text{i.e.} \quad = \quad M^*.$$

(vi) P.A.M. Dirac took the classical-radius of the electron e^2/m_ec^2 as a natural unit of length; and found an interesting relation:

$$R_0/r_e = e^2/Gm_em_p = \left[M_0/m_p\right]^{1/2} = 10^{40}.$$

Tank [4] explained the above relation and reached a conclusion that the relation implies: (i) Gravitational potential-energy of the universe is equal to the energy-of-mass of the universe; (ii) Electrostatic potential-energy of the electron is equal to the energy-of-mass of it; and (iii) Strengths of electric-force, strong-force and gravitational-force are proportional to densities of matter within the electron, the pi-meson and the universe respectively. Sidharth [5] and Funkhouser [6] have given a similar explanation for the recurrences of the Large-Number, but they have not drawn any conclusions for further application.

From the above comparison of Planck's natural units and Dirac's natural units we are led to a conclusion that Dirac's choice of natural units leads to interesting new relations. These relations should not be ignored as mere coincidences, because these relations have emerged from right choice of natural-units.

Sivaram [7] noticed the recurrences of the same value of acceleration, equal to the "critical-acceleration" of MOND, at the radial-distance R in the case of the electron, the proton, the nucleus, the globular-clusters, the spiral-galaxies, the galactic-clusters and the universe. Tank [8–10] could explain these recurrences based on equality of potential-energy and energy-of-mass of these systems, the equality which helped him to explain Dirac's large-number-ratios in 1997. Thus, Dirac's attempt to choose natural-units has led to a conclusion, of equality of potential-energy and energy-of-mass of various systems of matter, which helped explaining another set of recurrences noticed by Sivaram, and to draw further conclusions discussed in the references [8–10]

Also, if we measure distances in the units of radius-of-the-universe R_0 and measure masses of bodies in the units of total-mass-of-the-universe M_0 then the gravitational-constant G becomes unity; as follows:

> Gravitational-potential-energy of a system of masses M and m at a distance r is
>
> $$= (M/M_0)\left(mc^2\right)/(r/R_0).$$

Submitted on May 5, 2011 / Accepted on May 8, 2011

References

1. Ragazzoni R., Turatto M., Gaessler W. The lack of observational evidence for the quantum structure of space- time at Planck scales. *Astrophysical Journal*, 2003, v. 587, no. 2, L1–L4.
2. Lieu R., Hillman L. W. Phase coherence of light from extragalactic sources: Direct evidence against first-order Planck-scale fluctuations in "time" and "space". *Astrophysical Journal*, 2003, v. 585, L77–L80.
3. Whitfield J. Sharp images blur universal picture, *Nature*, 2003, DOI 10.1038/news030324-13, 31.
4. Tank H. K. Explanation for the recurrences of large-number 10^{40} in astrophysics and some insight into the strengths of fundamental forces in *Proceedings of Indian National Sci. Acad. A*, 1997, v. 63, no. 6, 469–474.
5. Sidharth B. The Planck Scale Underpinning for Spacetime. arXiv: gen-ph/0509026.
6. Funkhouser S. A New Large Number Coincidence and a Scaling Law for the Cosmological Constant. arXiv: gen-ph/0611115.
7. Sivaram C. Some aspects of MOND and its consequences for cosmology. *Astrophysics and Space Science*, 1994, v. 215, 185–189.
8. Tank H. K. Explanation for the recurrences of critical acceleration of MOND suggesting a clue to unification of fundamental forces. *Astrophysics and Space Science*, 2010, v. 330, 203–205.
9. Tank H. K. Explanation for the relative strengths of gravitational and electric forces. *Science and Culture*, 2009, v. 75, no. 9–10, 361–363.
10. Tank H. K. Wave-theoretical explanation for the newly-emerged-law of equality of potential-energy and energy-of-mass of reasonably independent systems of matter. *Advanced Studies in Theoretical Physics*, 2011, v. 5, no. 1, 45–55.

Excited Electronic States of Atoms Described by the Model of Oscillations in a Chain System

Andreas Ries and Marcus Vinicius Lia Fook

Universidade Federal de Campina Grande, UAEMA, Rua Aprígio Veloso 882, 58429-140 Campina Grande — PB, Brazil.
E-mail: andreasries@yahoo.com

We analyzed the numerical values of half-lifes of excited electronic states of the H, He and Li atom, as well as the Li^+ ion. By means of a fractal scaling model originally published by Müller in this journal, we interprete these half-lifes as proton resonance periods. On the logarithmic scale, the half-lifes were expressed by short continued fractions, where all numerators are Euler's number. From this representation it was concluded that the half-lifes are heavily located in nodes or sub-nodes of the spectrum of proton resonance periods.

1 Introduction

The model of a chain of similar harmonic oscillators was proposed by Müller [1–3] as a phenomenological theory describing physical quantities as proton resonance oscillation modes.

In the most general case, the spectrum of eigenfrequencies of a chain system of many proton harmonic oscillators is given by the continuous fraction equation [2]

$$f = f_p \exp S, \tag{1}$$

where f is any natural oscillation frequency of the chain system, f_p the oscillation frequency of one proton and S the continued fraction corresponding to f. S was suggested to be in the canonical form with all partial numerators equal 1 and the partial denominators are positive or negative integer values

$$S = n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \ldots}}}. \tag{2}$$

Particularly interesting properties arise when the numerator equals 2 and all denominators are divisible by 3. Such fractions divide the logarithmic scale in allowed values and empty gaps, i.e. ranges of numbers which cannot be expressed with this type of continued fractions. He showed that these continued fractions generate a self-similar and discrete spectrum of eigenvalues [1], that is also logarithmically invariant. Maximum spectral density areas arise when the free link n_0 and the partial denominators n_i are divisible by 3.

In two previous articles [4, 5] we applied a slightly modified model, where all numerators were substituted by Euler's number. This model was particularly successful describing specific features of the solar system [5].

However, the true physical meaning of the numerator e is not yet clear. It must now be investigated, for which type of data exactly this type of continued fractions can be applied. There might be some data sets, where the numerator is 2, as it was suggested by Müller in a patent [6].

In this article we analyzed a set of very accurately determined half-lifes of excited states of atoms on the logarithmic scale. We show that continued fractions with Euler's number as numerator are adequate to describe these data.

2 Data source and computational details

All atomic spectral data were taken from the web site of the National Institute of Standards and Technology (NIST) [7]. NIST maintains a critical selection of spectral data previously published in regular scientific journals. For the H, He and Li atom, reference was given to a publication by Wiese [8].

Table 1 shows such a data compilation for the Hydrogen atom. We consider here only experimentally observed emission lines (i.e. not Ritz lines), for which the transition probabilities have been determined. We numbered these lines in the order of increasing wavelength and eliminated lines with an already previously listed transition probability. For the Hydrogen atom, this procedure resulted then in a set of 109 lines which have all different transition probabilities. Also, if a transition probability has a numerical error higher than 1% (according to NIST), the corresponding line was ignored.

The transition probability as given by NIST has the unit of frequency $[s^{-1}]$ and is also called the Einstein A coefficient of spontaneous emission. Consider a large number of atoms in an excited state i, decaying to the ground state k (k could also be any lower lying excited state). Equation (3) is then the rate law

$$\frac{\partial N}{\partial t} = -A_{ik} N, \tag{3}$$

which results in

$$N(t) = N_0 \exp(-A_{ik} t), \tag{4}$$

where $N(t)$ is the number of excited atoms at time t, N_0 the number of excited atoms at $t = 0$ and A_{ik} the Einstein A coefficient for the transition $i \to k$. From this exponential law, the half-life $T_{1/2}$ of the transition $i \to k$ can be calculated as

$$T_{1/2} = \frac{\ln(2)}{A_{ik}}. \tag{5}$$

Table 1: Observed emission lines of the Hydrogen atom with corresponding wavelengths and transition probabilities. Obs.: Line no. 18 represents a forbidden transition.

Line no.	Wavelength [Å]	Transition probability [s^{-1}]	Line no.	Wavelength [Å]	Transition probability [s^{-1}]
1	918.125	5.0659×10^4	56	6562.72482	2.2448×10^7
2	919.342	7.8340×10^4	57	6562.77153	2.2449×10^7
3	920.947	1.2631×10^5	58	6562.79	4.4101×10^7
4	923.148	2.1425×10^5	59	6562.85175	6.4651×10^7
5	926.249	3.8694×10^5	60	8392.40	1.5167×10^3
6	930.751	7.5684×10^5	61	8413.32	1.9643×10^3
7	937.801	1.9728×10^7	62	8437.95	2.5804×10^3
8	937.814	1.6440×10^6	63	8467.26	3.4442×10^3
9	949.742	3.4375×10^7	64	8502.49	4.6801×10^3
10	949.742	4.1250×10^6	65	8545.38	6.4901×10^3
11	972.517	1.2785×10^7	66	8598.39	9.2117×10^3
12	972.541	6.8186×10^7	67	8665.02	1.3431×10^4
13	1025.728	1.6725×10^8	68	8750.46	2.0207×10^4
14	1025.728	5.5751×10^7	69	8862.89	3.1558×10^4
15	1215.6699	6.2648×10^8	70	9015.3	5.1558×10^4
16	1215.6699	6.2649×10^8	71	9229.7	8.9050×10^4
17	1215.6701	4.6986×10^8	72	9546.2	1.6506×10^5
18	1215.67312	2.495×10^{-6}	73	10049.8	3.3585×10^5
19	3656.65	9.9657×10^1	74	10938.17	7.7829×10^5
20	3657.25	1.1430×10^2	75	12818.072	2.2008×10^6
21	3658.04	1.3161×10^2	76	15560.46	3.6714×10^3
22	3658.65	1.5216×10^2	77	16411.36	1.6205×10^4
23	3659.41	1.7669×10^2	78	16811.10	2.5565×10^4
24	3660.32	2.0612×10^2	79	17366.885	4.2347×10^4
25	3661.27	2.4162×10^2	80	18179.21	7.4593×10^4
26	3662.22	2.8474×10^2	81	18751.3	8.9860×10^6
27	3663.41	3.3742×10^2	82	21661.178	3.0415×10^5
28	3664.65	4.0224×10^2	83	26258.71	7.7110×10^5
29	3666.08	4.8261×10^2	84	32969.8	6.9078×10^4
30	3667.73	5.8304×10^2	85	37405.76	1.3877×10^5
31	3669.45	7.0963×10^2	86	40522.79	2.6993×10^6
32	3671.32	8.7069×10^2	87	46537.8	3.2528×10^5
33	3673.81	1.0777×10^3	88	51286.5	3.6881×10^4
34	3676.376	1.3467×10^3	89	74599.0	1.0254×10^6
35	3679.370	1.7005×10^3	90	75024.4	1.5609×10^5
36	3682.823	2.1719×10^3	91	81548.4	3.3586×10^3
37	3686.831	2.8093×10^3	92	86644.60	5.0098×10^3
38	3691.551	3.6851×10^3	93	87600.64	3.9049×10^4
39	3697.157	4.9101×10^3	94	93920.3	7.8037×10^3
40	3703.859	6.6583×10^3	95	105035.07	1.2870×10^4
41	3711.978	9.2102×10^3	96	108035.9	2.2679×10^3
42	3721.946	1.3032×10^4	97	113086.81	8.2370×10^4
43	3734.369	1.8927×10^4	98	115395.4	3.3253×10^3
44	3750.151	2.8337×10^4	99	123719.12	4.5608×10^5
45	3770.633	4.3972×10^4	100	123871.53	2.3007×10^4
46	3797.909	7.1225×10^4	101	125870.5	5.0797×10^3
47	3835.397	1.2156×10^5	102	190619.6	2.2720×10^5
48	3889.064	2.2148×10^5	103	278035.0	1.2328×10^5
49	3970.075	4.3889×10^5	104	690717	2.7989×10^4
50	4101.734	9.7320×10^5	105	887610	1.8569×10^4
51	4340.472	2.5304×10^6	106	1118630	1.2709×10^4
52	4861.28694	9.6680×10^6	107	1387500	8.9344×10^3
53	4861.29776	9.6683×10^6	108	1694230	6.4283×10^3
54	4861.35	8.4193×10^6	109	3376000	2.0659×10^3
55	6562.70969	5.3877×10^7			

Finally, the numerical values of continued fractions were always calculated using the the Lenz algorithm as indicated in reference [9].

3 Results and discussion

Half-lifes of exited states of atoms are abundantly available from the NIST web site, however, only for the light atoms such as H, He and Li these data have a very high accuracy. Considering for instance Fe as a heavy element, most of the Einstein A coefficients have uncertainties of 10-18% and are consequently not suitable for a numerical analysis.

Due to results form our previous publications, we suspect that Müller's continued fraction formalism with Euler's number as numerator can still be applied to many data sets, so we set all partial numerators in Müller's continued fractions (given in equation (2)) to Euler's number.

We strictly follow the formalism of previous publications [4–6] and introduce a phase shift p in equation (2). According to [6] the phase shift can only have the values 0 or $\pm\frac{3}{2}$. So we write for the half-lifes of the excited states:

$$\ln \frac{T_{1/2}}{\tau} = p + S, \tag{6}$$

where S is the continued fraction

$$S = n_0 + \cfrac{e}{n_1 + \cfrac{e}{n_2 + \cfrac{e}{n_3 + ...}}} \tag{7}$$

and $\tau = \frac{\lambda_C}{c}$ is the oscillation period of a hypothetical photon with the reduced Compton wavelength of the proton ($\lambda_C = \frac{h}{2\pi mc} = 2.103089086 \times 10^{-16}$ m) and traveling at light speed (numerical value $7.015150081 \times 10^{-25}$ s).

We abbreviate $p + S$ as $[p; n_0 \mid n_1, n_2, n_3, \ldots]$. The free link n_0 and the partial denominators n_i are integers divisible by 3. For convergence reason, we have to include $|e{+}1|$ as allowed partial denominator. This means the free link n_0 is allowed to be $0, \pm3, \pm6, \pm9 \ldots$ and all partial denominators n_i can take the values $e{+}1$, $-e{-}1$, $\pm6, \pm9, \pm12 \ldots$.

For the calculation of the continued fractions we did not consider any standard deviation of the published data. Practically, we developed the continued fraction and determined only 18 partial denominators. Next we calculated repeatedly the data value from the continued fraction, every time considering one more partial denominator. As soon as considering further denominators did not improve the experimental data value significantly (on the linear scale), we stopped considering further denominators and gave the resulting fraction in Table 2. This means we demonstrate how accurately the calculated half-lifes can be expressed through continued fractions. Additionally we also report the numerical error, which is defined as absolute value of the difference between the half-life calculated from the NIST transition probability and the value calculated from the continued fraction representation.

If this numerical error is higher than 1%, we interpret the result as "no continued fraction found", otherwise the continued fraction representation is in satisfying agreement with the experimental data.

As can be seen from Table 2, with one exception, all half-lifes could be expressed in a satisfactory manner by a continued fraction representation. Only one outlier was found,

which underlines the statistical nature of Müller's continued fraction model.

We believe that spectral line number 71 is a true outlier rather than a bad data point, since the Hydrogen spectrum has been thoroughly investigated and is definitely the most easiest one to interpret.

In most cases the numerical errors are several orders of magnitude lower than the data value. This changes when calculating the continued fractions with number 2 as numerator, as it was suggested by Müller in a patent [6]. In this case the number of outliers increases to 12 and the numerical errors of the continued fraction representations are frequently very slightly lower than the 1% limit. So the numerator e is definitely the better choice.

It can be seen that around 25% of the half-lifes could be expressed by two continued fractions, so there is no preferred accumulation of the half-lifes in the neutral zones. The majority of the continued fraction representations terminates with a high partial denominator ($\pm 9, \pm 12, \pm 15 \ldots$). This means there is a general tendency that the half-lifes accumulate in nodes and sub-nodes of the spectrum of the proton resonance periods.

Additionally, in the same manner as here described for the spectral lines of the Hydrogen atom, we analyzed the spectral data of He, Li (neutral atoms) and the Li^+ ion. From the NIST database resulted 142 spectral lines for the He atom, 57 lines for the Li atom and 129 lines for the Li^+ ion.

Again, it was analogously possible to express the half-lifes on the logarithmic scale by continued fraction representations with Euler's number as numerators. Very few outliers were found, 6 in the He data set, only one in the Li data set and 7 in the set of the Li^+ lines (continued fractions not given). Regarding the numerical errors, no significant differences were detected, when comparing with the Hydrogen set.

This result is a contribution to the importance of Euler's number as a possible numerator in the model of oscillations in a chain system. We have now identified the half-lifes of excited states with respect to individual electronic transitions as a further data set where this (still phenomenological) model can be applied. For the half-lifes, apparently it does not matter how many nucleons are in the atom and whether the atom is neutral or charged. It even seems to be that the model applies for both, allowed and forbidden transitions, however, this should be verified with further data; we have here only one forbidden transition in our data set.

4 Conclusions

Numerical investigation of a large data set of 437 half-lifes of electronic transitions from different atoms revealed that Müller's continued fraction model with e as numerator is adequate to express these data on the logarithmic scale. There is a general tendency that half-lifes accumulate in nodes and sub-nodes of the spectrum of proton resonance periods. This accumulation does not seem to be influenced by the atomic charge or the atomic number (chemical element). It can be said that every excited state of an atom (with corresponding transition), has different oscillation properties and goes in resonance with the appropriate proton oscillation. Then, during one proton oscillation period, 50% of the excited atoms become de-excited to a lower-lying state.

This viewpoint has some similarity to the teaching of modern quantum electrodynamics. This theory states that spontaneous emission from atoms is caused by a 50:50 contribution from radiation reaction and vacuum fluctuations [10]. So both models assume an external influence coupled to the atoms, either the proton resonance spectrum or the vacuum fluctuations.

Acknowledgments

The authors greatly acknowledge the financial support from the Brazilian governmental funding agencies FAPESQ and CNPq.

Submitted on June 7, 2011 / Accepted on June 9, 2011

References

1. Müller H. Fractal scaling Models of resonant oscillations in chain systems of harmonic oscillators. *Progress in Physics*, 2009, v. 2 72–76.
2. Müller H. Fractal scaling models of natural oscillations in chain systems and the mass distribution of the celestial bodies in the solar system. *Progress in Physics*, 2010, v. 1 62–66.
3. Müller H. Fractal scaling models of natural oscillations in chain systems and the mass distribution of particles. *Progress in Physics*, 2010, v. 3 61–66.
4. Ries A., Fook M. V. L. Fractal structure of nature's preferred masses: Application of the model of oscillations in a chain system. *Progress in Physics*, 2010, v. 4, 82–89.
5. Ries A., Fook M. V. L. Application of the model of oscillations in a chain system to the solar system. *Progress in Physics*, 2011, v. 1, 103–111.
6. Otte R., Müller H. German patent No. DE102004003753A1, date: 11.08.2005
7. NIST ASD Team (2010). NIST Atomic Spectra Database (ver. 4.0.1), [Online]. Available: http://physics.nist.gov/asd [2011, May 23]. National Institute of Standards and Technology, Gaithersburg, MD.
8. Wiese W. L., Fuhr J. R. Accurate atomic transition probabilities for hydrogen, helium, and lithium. *Journal of Physical and Chemical Reference Data*, 2009, v. 38, 565–726, Erratum: 2009, v. 38, 1129.
9. Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P. Numerical recipes in C. Cambridge University Press, Cambridge, 1992.
10. Milonni P. W. The quantum vacuum. An introduction to quantum electrodynamics. Academic Press, San Diego, 1993.

Table 2: Continued fraction representation of half-lifes of excited states of the Hydrogen atom

Line no.	Half-life [s] Continued fraction representation	Numerical error [s]
1	$1.36826068529 \times 10^{-5}$	
	[0; 45 | -e-1, -e-1, e+1, -6, 6, -e-1, 6]	2.5×10^{-11}
	[1.5; 42 | e+1, -e-1, 24, -6, 9]	1.5×10^{-11}
2	$8.84793439571 \times 10^{-6}$	
	[1.5; 42 | 6, -9, e+1]	3.7×10^{-8}
3	$5.48766669749 \times 10^{-6}$	
	[1.5; 42 | 765]	3.0×10^{-11}
4	$3.23522604695 \times 10^{-6}$	
	[0; 42 | e+1, -e-1, e+1, -9, 12]	4.2×10^{-11}
	[1.5; 42 | -6, e+1, -6, -e-1, -9]	1.6×10^{-11}
5	$1.79135571551 \times 10^{-6}$	
	[0; 42 | 6, e+1, -e-1]	1.7×10^{-8}
6	$9.15843745785 \times 10^{-7}$	
	[0; 42 | -9, -6, 9, 6, -21, 117]	1.7×10^{-19}
7	$3.51351977169 \times 10^{-8}$	
	[0; 39 | -6, e+1, -e-1]	4.7×10^{-10}
	[1.5; 36 | e+1, -e-1, e+1, e+1, -e-1]	4.4×10^{-11}
8	$4.21622372603 \times 10^{-7}$	
	[1.5; 39 | 6, 12, e+1, 6]	6.2×10^{-12}
9	$2.01642816163 \times 10^{-8}$	
	[1.5; 36 | 6, e+1, -6, 6, -e-1, 18]	2.0×10^{-15}
10	$1.68035680136 \times 10^{-7}$	
	[1.5; 39 | -6, 6, e+1, -e-1, e+1]	3.4×10^{-10}
11	$5.42156574548 \times 10^{-8}$	
	[0; 39 | -24, 27, -24, -18]	3.6×10^{-17}
12	$1.01655351621 \times 10^{-8}$	
	[1.5; 36 | -9, -6, -42]	4.2×10^{-14}
13	$4.1443777612 \times 10^{-9}$	
	[0; 36 | 9, -6, -e-1]	4.8×10^{-12}
14	$1.24329102717 \times 10^{-8}$	
	[1.5; 36 | -33]	4.9×10^{-11}
15	$1.106415497 \times 10^{-9}$	
	[1.5; 33 | 6, -6, e+1, e+1, -e-1, 12]	1.1×10^{-15}
16	$1.10639783645 \times 10^{-9}$	
	[1.5; 33 | 6, -6, e+1, e+1, -e-1, 6, 15]	4.3×10^{-17}
17	$1.47522066267 \times 10^{-9}$	
	[0; 36 | -e-1, -39, -e-1, e+1, 9, 6]	4.8×10^{-17}
	[1.5; 33 | e+1, -12, e+1, -9, 12, 12]	8.9×10^{-18}
18	277814.501226	
	[0; 69 | -e-1, 6, -e-1, -141]	0.17
	[1.5; 66 | e+1, 6, -81]	0.97
19	0.00695532858264	
	[0; 51 | -9, e+1, -e-1, e+1]	9.2×10^{-5}
20	0.00606427979493	
	[0; 51 | -6, 6, 15]	5.7×10^{-7}
21	0.00526667563681	
	[0; 51 | -e-1, -e-1, -e-1, 6, 12]	4.2×10^{-8}
	[1.5; 48 | e+1, -e-1, -e-1, 6, -e-1, e+1, -30]	6.8×10^{-10}
22	0.00455538367876	
	[0; 51 | -e-1, 12, e+1, -e-1, 12, 9]	8.9×10^{-10}
	[1.5; 48 | e+1, 90, -e-1, e+1, 9]	7.1×10^{-10}
23	0.00392295648062	
	[0; 51 | -e-1, e+1, -12, -12, 54]	1.2×10^{-10}
24	0.00336283320668	
	[1.5; 48 | 6, 6, 39, -6, -18]	2.1×10^{-11}
25	0.00286874919527	
	[1.5; 48 | 9, e+1, -e-1, e+1, -e-1]	2.7×10^{-5}
26	0.00243431615003	
	[1.5; 48 | 27, e+1, e+1, -18]	3.2×10^{-9}
27	0.00205425635872	
	[1.5; 48 | -39, e+1, -e-1, -e-1, e+1, -33]	2.2×10^{-11}
28	0.00172321793099	
	[1.5; 48 | -12, e+1, -e-1, e+1, -e-1, e+1, 12]	2.3×10^{-9}
29	0.00143624703293	
	[1.5; 48 | -6, -9, e+1, -6, e+1]	3.7×10^{-8}
30	0.00118885013131	
	[0; 48 | e+1, -e-1, -e-1, -e-1, e+1]	1.4×10^{-6}
	[1.5; 48 | -e-1, -e-1, -9, -6, 9]	1.6×10^{-9}
31	0.000976772656962	
	[0; 48 | e+1, 12, -e-1, e+1]	5.4×10^{-7}
32	0.000796089515855	
	[0; 48 | 6, -9, e+1, -e-1]	6.6×10^{-7}
33	0.000643172664526	
	[0; 48 | 9, e+1, -e-1]	4.4×10^{-6}
34	0.000514700512779	
	[0; 48 | 60, e+1, -e-1, -60]	2.0×10^{-11}
35	0.000407613749227	
	[0; 48 | -15, e+1, e+1, -e-1, -9]	1.8×10^{-9}
36	0.000319143229688	
	[0; 48 | -6, -9, -e-1, e+1, -6]	4.3×10^{-9}
37	0.000246733058256	
	[0; 48 | -e-1, -12, -6, 6]	3.3×10^{-9}
38	0.000188094537614	
	[0; 48 | -e-1, e+1, -e-1, -9, 6]	6.2×10^{-9}
	[1.5; 45 | 6, -e-1, e+1, -6, e+1, -351]	6.9×10^{-13}
39	0.000141167630101	
	[1.5; 45 | 12, -e-1, e+1]	8.2×10^{-7}
40	0.000104102726005	
	[1.5; 45 | -51, 9]	4.2×10^{-9}
41	$7.525864591 \times 10^{-5}$	
	[1.5; 45 | -6, -e-1, e+1, -e-1]	9.0×10^{-7}
42	$5.31880893616 \times 10^{-5}$	
	[0; 45 | e+1, -12, -e-1, e+1, -9]	5.3×10^{-10}
	[1.5; 45 | -e-1, -90, e+1, 60]	5.1×10^{-13}
43	$3.66221366598 \times 10^{-5}$	
	[0; 45 | 6, e+1, -15, -e-1, -e-1, 18]	5.8×10^{-13}
44	$2.44608526153 \times 10^{-5}$	
	[0; 45 | -1446]	3.6×10^{-11}
45	$1.57633762522 \times 10^{-5}$	
	[0; 45 | -6, -18, e+1]	3.1×10^{-9}
46	$9.73179614686 \times 10^{-6}$	
	[0; 45 | -e-1, e+1, -12, -e-1, e+1, 9]	1.8×10^{-11}
	[1.5; 42 | e+1, e+1, -e-1, e+1, -e-1, e+1, -6, 6]	2.0×10^{-11}
47	$5.70209921487 \times 10^{-6}$	
	[1.5; 42 | 66, -e-1, e+1, -e-1]	4.6×10^{-10}
48	$3.12961522738 \times 10^{-6}$	
	[0; 42 | e+1, -e-1, 6, -31650]	7.5×10^{-17}
49	$1.57931869161 \times 10^{-6}$	
	[0; 42 | 12, -e-1, e+1, -e-1]	1.8×10^{-8}
50	$7.12235080723 \times 10^{-7}$	
	[0; 42 | -6, e+1, -e-1, 6]	6.3×10^{-10}
	[1.5; 39 | e+1, -e-1, e+1, 9]	4.4×10^{-10}
51	$2.73927908852 \times 10^{-7}$	
	[1.5; 39 | 441, e+1, 12]	1.6×10^{-14}
52	$7.16949917832 \times 10^{-8}$	
	[0; 39 | 15, e+1, -e-1]	3.7×10^{-10}
53	$7.16927671421 \times 10^{-8}$	
	[0; 39 | 15, e+1, -e-1]	3.7×10^{-10}
54	$8.23283622819 \times 10^{-8}$	
	[0; 39 | 9, -48]	5.0×10^{-12}
55	$1.2865363338 \times 10^{-8}$	
	[1.5; 36 | -51, -e-1]	4.8×10^{-12}
56	$3.08779036244 \times 10^{-8}$	
	[0; 39 | -e-1, -9, -30]	8.1×10^{-13}
	[1.5; 36 | e+1, -6, -6, e+1, -9]	4.1×10^{-13}

Line no.	Half-life [s] Continued fraction representation	Numerical error [s]
57	$3.08765281554 \times 10^{-8}$	
	[0; 39 | -e-1, -9, -27, e+1, 30]	5.5×10^{-16}
	[1.5; 36 | e+1, -6, -6, e+1, -6, -69]	2.4×10^{-16}
58	$1.57172667413 \times 10^{-8}$	
	[1.5; 36 | 18, 9]	7.5×10^{-12}
59	$1.07213682783 \times 10^{-8}$	
	[1.5; 36 | -12, 6, e+1, 9]	8.2×10^{-14}
60	0.000457010074873	
	[0; 48 | -36, -e-1, -e-1, 72]	8.3×10^{-12}
61	0.000352872361941	
	[0; 48 | -9, e+1, -6, 66]	6.5×10^{-11}
62	0.000268620051372	
	[0; 48 | -e-1, -e-1, 15, -6, 6]	8.3×10^{-10}
	[1.5; 45 | e+1, -e-1, -9, -6, e+1, 6, -135]	3.4×10^{-14}
63	0.000201250560525	
	[0; 48 | -e-1, e+1, 9, 6]	2.9×10^{-8}
	[1.5; 45 | e+1, e+1, -15, e+1, 6, e+1, -72]	4.0×10^{-14}
64	0.000148105207273	
	[1.5; 45 | 9, 30, -6, e+1, -6, e+1, e+1]	7.5×10^{-13}
65	0.00010680069345	
	[1.5; 45 | -96, -e-1, e+1]	6.3×10^{-9}
66	$7.52463910635 \times 10^{-5}$	
	[1.5; 45 | -6, -e-1, e+1, -e-1]	8.9×10^{-7}
67	$5.16080098697 \times 10^{-5}$	
	[0; 45 | e+1, -39, -e-1, 6, -e-1, -9, -6, 18]	1.6×10^{-15}
	[1.5; 45 | -e-1, 24]	6.5×10^{-8}
68	$3.43023299134 \times 10^{-5}$	
	[0; 45 | 9, -e-1, e+1, -27, -6]	3.1×10^{-11}
69	$2.19642303238 \times 10^{-5}$	
	[0; 45 | -24, -e-1, 6, e+1]	4.6×10^{-10}
70	$1.34440277078 \times 10^{-5}$	
	[0; 45 | -e-1, -e-1, 9, -e-1, e+1]	9.5×10^{-9}
	[1.5; 42 | e+1, -e-1, -15, 12, -e-1, -15]	6.3×10^{-13}
71	$7.78379764806 \times 10^{-6}$	
	[1.5; 42 | 9, -e-1, e+1, -e-1, e+1, -e-1, e+1, -e-1]	1.0×10^{-7}
	no continued fraction found	error 1.3%
72	$4.19936496159 \times 10^{-6}$	
	[1.5; 42 | -9, -e-1, e+1, -e-1, e+1, -e-1]	3.3×10^{-8}
73	$2.06385940319 \times 10^{-6}$	
	[0; 42 | 6, -e-1, 6, 12, 6, 12]	1.3×10^{-13}
	[1.5; 42 | -e-1, e+1, -e-1, 9, e+1, 12]	4.2×10^{-12}
74	$8.90602706652 \times 10^{-7}$	
	[0; 42 | -9, 6, e+1, -e-1]	9.1×10^{-10}
75	$3.14952372119 \times 10^{-7}$	
	[1.5; 39 | 18, e+1, 6, 9, -e-1, -6, 9, -9]	2.6×10^{-17}
76	0.000188796421136	
	[0; 48 | -e-1, e+1, -e-1, -e-1, -e-1, e+1]	7.0×10^{-8}
	[1.5; 45 | 6, -e-1, e+1, -e-1, 6]	6.1×10^{-8}
77	$4.27736612502 \times 10^{-5}$	
	[1.5; 45 | -e-1, e+1, -6, 15]	4.4×10^{-9}
78	$2.71131304737 \times 10^{-5}$	
	[0; 45 | 27, -27, e+1, e+1, 9]	2.0×10^{-13}
79	$1.63682712013 \times 10^{-5}$	
	[0; 45 | -6, -e-1, 138]	3.9×10^{-11}
80	$9.29238910568 \times 10^{-6}$	
	[0; 45 | -e-1, e+1, -e-1, 33]	6.2×10^{-10}
	[1.5; 42 | 6, -e-1, e+1, e+1, -15]	1.3×10^{-10}
81	$7.71363432628 \times 10^{-8}$	
	[0; 39 | 12, -e-1, -e-1, -9]	2.2×10^{-12}
82	$2.27896492047 \times 10^{-6}$	
	[0; 42 | e+1, e+1, e+1, e+1, -e-1, e+1, -6, -15]	1.3×10^{-13}
	[1.5; 42 | -e-1, e+1, e+1, 45]	2.9×10^{-11}

Line no.	Half-life [s] Continued fraction representation	Numerical error [s]
83	$8.9890699074 \times 10^{-7}$	
	[0; 42 | -9, 27]	7.5×10^{-11}
84	$1.00342682266 \times 10^{-5}$	
	[0; 45 | -e-1, e+1, 9, -9]	7.6×10^{-10}
	[1.5; 42 | e+1, e+1, -21]	2.9×10^{-9}
85	$4.9949353647 \times 10^{-6}$	
	[1.5; 42 | -30, -105, 6]	3.2×10^{-13}
86	$2.56787752588 \times 10^{-7}$	
	[1.5; 39 | -48, e+1]	2.5×10^{-10}
87	$2.130924682 \times 10^{-6}$	
	[1.5; 42 | -e-1, e+1, -6, 39, -30]	8.3×10^{-14}
88	$1.87941536444 \times 10^{-5}$	
	[0; 45 | -9, -e-1, e+1, -e-1]	1.4×10^{-7}
89	$6.75977355725 \times 10^{-7}$	
	[0; 42 | -e-1, -e-1, e+1, 6, -6, 12]	8.3×10^{-13}
	[1.5; 39 | e+1, -e-1, -312, 24]	5.5×10^{-15}
90	$4.44068922135 \times 10^{-6}$	
	[1.5; 42 | -12, -e-1, e+1, -e-1]	6.5×10^{-9}
91	0.000206379795319	
	[0; 48 | -e-1, e+1, e+1, -e-1, -e-1, -e-1, e+1, 909]	1.2×10^{-14}
	[1.5; 45 | e+1, e+1, e+1, -e-1, -e-1, -e-1, -12]	5.0×10^{-11}
92	0.000138358253934	
	[1.5; 45 | 12, -12, 111]	1.5×10^{-11}
93	$1.77507024651 \times 10^{-5}$	
	[0; 45 | -9, e+1, e+1]	2.7×10^{-8}
94	$8.88228892141 \times 10^{-5}$	
	[1.5; 45 | -12, -e-1, 9, -e-1, e+1, e+1, -6, -6]	4.5×10^{-13}
95	$5.38575897871 \times 10^{-5}$	
	[0; 45 | e+1, -9, -e-1, e+1, -e-1]	8.3×10^{-8}
	[1.5; 45 | -e-1, -27]	4.5×10^{-8}
96	0.000305633925905	
	[0; 48 | -6, 9, 15, 21, -e-1, -12, -e-1, -18]	1.9×10^{-16}
97	$8.41504407624 \times 10^{-6}$	
	[1.5; 42 | 6, 9, -24, 6, -33, -e-1, 12]	7.3×10^{-17}
98	0.000208446510258	
	[0; 48 | -e-1, 6, -e-1, e+1, -e-1]	8.3×10^{-7}
	[1.5; 45 | e+1, 6, -e-1, 6]	5.7×10^{-8}
99	$1.51979297614 \times 10^{-6}$	
	[0; 42 | 12, 6, e+1]	7.1×10^{-10}
100	$3.01276646481 \times 10^{-5}$	
	[0; 45 | 12, e+1, -e-1, e+1]	9.4×10^{-8}
101	0.000136454353714	
	[1.5; 45 | 12, 6, -e-1, 6, e+1, -6, -12]	1.6×10^{-12}
102	$3.0508238581 \times 10^{-6}$	
	[0; 42 | e+1, -e-1, 27]	7.5×10^{-10}
	[1.5; 42 | -e-1, -e-1, e+1, -6, -9]	8.1×10^{-11}
103	$5.62254364504 \times 10^{-6}$	
	[1.5; 42 | 99, -e-1, e+1]	6.6×10^{-10}
104	$2.476498555 \times 10^{-5}$	
	[0; 45 | 261, -e-1]	7.8×10^{-10}
105	$3.73281911013 \times 10^{-5}$	
	[0; 45 | 6, 6, -30, -6, -6]	2.1×10^{-12}
106	$5.4539678543 \times 10^{-5}$	
	[0; 45 | e+1, -9, 6, -e-1, -30]	4.0×10^{-11}
	[1.5; 45 | -e-1, -18, e+1]	7.3×10^{-8}
107	$7.75818387983 \times 10^{-5}$	
	[1.5; 45 | -9, e+1, -e-1, e+1, -e-1]	6.5×10^{-7}
108	0.000107827447468	
	[1.5; 45 | -147, -6, e+1]	4.3×10^{-10}
109	0.000335518263498	
	[0; 48 | -6, -e-1, e+1, -e-1, e+1]	1.9×10^{-6}

Phenomenological Derivation of the Schrödinger Equation

Fernando Ogiba

91220-480, Porto Alegre, RS, Brazil
E-mail: ogiba@cpovo.net

The Schrödinger equation is derived classically assuming that particles present local random spatial fluctuations compatible with the presence of the zero-point field. Without specifying the forces arising from this permanent matter-field interaction but exploring its fundamental properties (homogeneity, isotropy and random aspect) to justify the emergence of the continuity equation in one-particle context, these fluctuations are described in terms of the probability density. Specifically, the starting point is the assumption that the local activities, which turn the path followed by the particle totally unpredictable, must be associated with an energy proportional to $\partial P/\partial t$. The polar form of the wave function, which connects the obtained classical equations with the corresponding quantum equation, emerges as a by-product of the approach.

1 Introduction

The evolution of the wave function in single-particle quantum systems is described by the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t}, \tag{1}$$

where m is the mass and V is a potential. The complex wave function is generally presented in its polar form

$$\psi = \sqrt{P}\exp(iS/\hbar), \tag{2}$$

where $P=|\psi|^2$ is the probability density, and $S/\hbar$ is a phase. Substituting (2) into (1) results in two equations

$$\frac{\partial P}{\partial t} + \nabla\cdot\left(P\frac{\nabla S}{m}\right) = 0, \tag{3}$$

and

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V + Q = 0, \tag{4}$$

where

$$Q = -\frac{\hbar^2}{4m}\left[\frac{\nabla^2 P}{P} - \frac{1}{2}\frac{(\nabla P)^2}{P^2}\right] \tag{5}$$

is known as quantum potential. At the classical limit ($\hbar \to 0$) Q vanishes and (4) reduces to the Hamilton-Jacobi equation. For this reason, Bohm [1] suggested that S is the classical action function, which relates to the actual velocity, $\mathbf{v} = \nabla S/m$, of the particle. In this way (3) simply expresses the conservation of probability.

This alternative way of writing the Schrödinger equation presents advantages as regards its interpretation in terms of classical variables. However, the problem of ignoring the path followed by the particle persists. And more, we have an obvious increase in complexity: The Schrödinger equation is a single function and quite simple, on the other hand, the equation (4) is somewhat complicated - and still requires the continuity equation to account local activities. And above all, thinking that the quantum revolution, highly non-classical, has its origin in a classical equation with an additional potential is not very easy. In reality, Q is not a traditional potential, but part of the description of the motion, that is, P is playing the role of a dynamical variable at the same footing as S. Thus S and P can be said to codetermine each other. However, in approximate schemes to get information about quantum systems it can be used as a potential [2].

Equation (4) is referred as stochastic Hamilton-Jacobi-Bohm equation. Despite the fact that P is unique for a given quantum system, it is interpreted as a differential equation describing an ensemble of trajectories. This is grounded in the fact that the action S was originally defined as a field variable related with a set of potential trajectories [3].

It is paid much attention to equation (4) and less concern about (3). From a dynamical point of view, the emergence of the continuity equation is the most remarkable result: It highlights the local loss of determinism ($\partial P/\partial t \neq 0$), is valid for one-particle systems (it was obtained in this way), and contains inherently the multiple path aspect of quantum systems [4], exactly how is assigned to equation (4).

Fundamentally, to have $\partial P/\partial t \neq 0$ (change of probability at a given position), and thus to justify the emergence of the equation (3), it is necessary that the particle runs local random spatial fluctuations. Otherwise, there are local preferences, and these combined with the dynamics that emanates from the potential V (deterministic) results in a classical trajectory. Therefore these fluctuations require the presence of external forces with special features. Indeed, these forces exist and are related with the zero-point field (ZPF). They are formally treated in the context of the stochastic electrodynamics [5, 6], and under certain conditions they may be measured [7, 8]. However, their definition is outside the scope of this work; just let's enumerate its indispensable characteristics to justify

the continuity equation in the context of one-particle dynamics.

The above rewriting of the Schrödinger equation starts from propositions valid within the quantum formalism and arrives at seemingly classical equations. What will be done in the present paper is to follow an inverse path. The starting point is the fact that the local changes of the probability density — associated with isotropic random spatial fluctuations impressed by the ZPF — must be related with an energy.

2 The multi path aspect of the motion

Suppose a particle of mass m performing a motion with velocity $\mathbf{v}$. If the associated probability density P is a continuous function of the coordinates and time, then its dynamical evolution along the trajectory is given by

$$\frac{dP}{dt} = \frac{\partial P}{\partial t} + \mathbf{v} \cdot \nabla P, \tag{6}$$

where $\partial P/\partial t$ refers to the change of probability at a given position, and the second term accounts for the spatial changes. As P is a probability, then we cannot precise the angle between ∇P and $\mathbf{v}$. Moreover, in principle, ∇P can show an isotropic distribution around each position. Indeed, as P is a conserved quantity, then the change of the probability density inside a given volume Ω (arbitrary), containing the instantaneous position of the particle, must be equal to the probability flux through a surface A surrounding this volume. Formally, we have

$$\frac{\partial}{\partial t}\int_{\Omega} P d\Omega = -\int_{A} P\dot{\mathbf{r}} \cdot d\mathbf{A}, \tag{7}$$

where $\dot{\mathbf{r}}$ is a velocity, and the vector field $P\dot{\mathbf{r}}$ represents all possible probability currents that cross the surface A. Obviously, if the particle is inside this volume, it emerges following one of these possibilities. In accordance with the properties of the ZPF, the field $P\dot{\mathbf{r}}$ must present an isotropic distribution, however, as the velocity of the particle is dictated by the dynamics of the system as a whole, then there are some privileged probability currents (the resulting motion is not a random walk). According to Green's theorem and equation (7), each one of the possible currents obeys

$$\frac{\partial P}{\partial t} + \nabla \cdot (P\dot{\mathbf{r}}) = 0. \tag{8}$$

As this process is repeated at all positions where the particle can be found, linking the successive probability currents, according to which the particle emerges from each volume Ω, is defined a path described by the velocity

$$\mathbf{v} = \frac{\nabla S}{m}, \tag{9}$$

where S is the Hamilton-Jacobi function of one possible path [3, see p. 36]. Therefore equation (8) must be written as (3).

If the local activities are ignored (classical limit), then the function S is defined on a single trajectory. This also can be easily inferred making $\partial P/\partial t = 0$ in equation (7). In this case the probability flux that enters the volume Ω equals the one that emerges from it. This means that the particle has only one possibility (probability current) to leave each successive volume Ω.

If the external field acts on the particle everywhere (homogeneously), without preferred directions (isotropic) and in a totally unpredictable (random) way, that is, like the ZPF, then we will have a local motion compatible with the continuity equation. Therefore, as the particle has several possibilities to leave each position (following one possible current $P\dot{\mathbf{r}}$), this assigns a multi path aspect to the motion. This means that the particle can travel on each one of them indiscriminately; there is no preferred path. Note, not having a preferred path means that all are equally probable. We realize that this fact is consistent with the formulation of quantum mechanics in terms of path integrals, where Feynman and Hibbs [4, see p. 28] begin with the following statement: "Now we can give the quantum-mechanical rule. We must say how much each trajectory contributes to the total amplitude to go from a to b. It is not that just the particular path of extreme action contributes; rather, it is that all the paths contribute. They contribute in equal amounts to the total amplitude, but contribute as different phases. The phase of the contribution from a given path is the action S for that path in units of the quantum of action $\hbar$". Coincidently, this is a description of the evolution operator $\exp(iS/\hbar)$ (unitary), present in (2), which is the core of the path integrals.

3 The main proposition

In a classical system, the particles are actuated by forces in such a way that they move along single predictable trajectories, and this leads to $\partial P/\partial t = 0$ everywhere (the local activities are ignored). By other side, if particles are being actuated by a field, with the characteristics pointed above, local exchange of energy between them occurs in such a way that $\partial P/\partial t \neq 0$. Admitting that this is a fact, let's write an effective stationary action function S_{eff} that, in addition to describing a path through the function S, also takes into account the local activities described in terms of probability density, that is,

$$S_{eff} = S + S_l, \tag{10}$$

where S_l is a local action that depends only on P. Following the same formalism obeyed by the stationary Hamilton's function, the energy and momentum of the particle over a possible path are, respectively, written as

$$H = -\frac{\partial S_{eff}}{\partial t} = -\frac{\partial S}{\partial t} - F\frac{\partial P}{\partial t}, \tag{11}$$

and

$$\mathbf{p} = \nabla S_{eff} = \nabla S + F\nabla P, \tag{12}$$

where $F = \partial S_l/\partial P$ should be a function of P which must comply the dynamics of the system. Specifically, this function must obey the conservation of probability and the local conservation of energy (the particle cannot extract energy from the field indefinitely).

The motion equations of the system can be obtained in the following way: As S and P are taking values on a volume, then the average energy of the multi path system need to be written in the form

$$\bar{H} = \int d^3rPH = \int d^3r\mathcal{H}, \tag{13}$$

where the integral is taken over whole space. Here, $\mathcal{H}$ has the role of Hamiltonian density. With H given by (11) we have

$$\bar{H} = \int d^3rP\left(-\frac{\partial S}{\partial t} - F\frac{\partial P}{\partial t}\right). \tag{14}$$

As $\bar{H}$, written in this way, is a functional of the functions S and P, taking the functional derivatives with respect to these functions, according to the well known rules

$$\frac{\delta\bar{H}}{\delta\xi} = \frac{\partial\mathcal{H}}{\partial\xi} - \frac{\partial}{\partial x_\alpha}\left(\frac{\partial\mathcal{H}}{\partial(\partial\xi/\partial x_\alpha)}\right), \tag{15}$$

where $x_\alpha = x, y, z, t$ and $\xi = S$ or P, we obtain respectively

$$\frac{\delta\bar{H}}{\delta S} = \frac{\partial P}{\partial t} \tag{16}$$

and

$$\frac{\delta\bar{H}}{\delta P} = -\frac{\partial S}{\partial t}. \tag{17}$$

This shows that the proposition (10) preserves the shapes of the canonical equations, where S and P behave as dynamical conjugate variables of the canonically transformed Hamiltonian $\bar{H}$ [1].

Taking into account the momentum (12), the energy (11) can be expressed by

$$H = \frac{|\nabla S + F\nabla P|^2}{2m} + V, \tag{18}$$

then (13) can also be written as

$$\bar{H} = \int d^3rP\left(\frac{|\nabla S + F\nabla P|^2}{2m} + V\right), \tag{19}$$

and, consequently, the canonical equation (16) takes the form

$$\left[\frac{\partial P}{\partial t} + \nabla\cdot\left(P\frac{\nabla S}{m}\right)\right] + (F + PF')\frac{(\nabla P)^2}{m} + PF\frac{\nabla^2 P}{m} = 0, \tag{20}$$

where $F' = \partial F\,/\,\partial P$. The first term, being the continuity equation, is zero, and the trivial solution of the resulting equation gives simultaneously $F = cte/P$ and $F = 0$. However, if this trivial solution is valid, F is not defined in the field of real numbers.

Generalizing the constant to complex numbers, the non zero solution is written as $F = (S_1 + iS_0)/P$, where S_1 and S_0 are real constants (they have dimension of action). Thus, returning this complex shape of F into (19), from (16), results

$$\left[\frac{\partial P}{\partial t} + \nabla\cdot\left(P\frac{\nabla S}{m}\right)\right] + S_1\frac{\nabla^2 P}{m} = 0, \tag{21}$$

which shows that probability conservation is obeyed if F is a pure imaginary ($S_1 = 0$). As this occurs independently of the P^{-1} functionality, then it only justifies the complex aspect of the trivial solution of (20).

Another evidence that F is pure imaginary comes from the fact that the momentum (12) is apparently incompatible with the actual velocity (9); it seems that we should have

$$\mathbf{v} = \frac{\nabla S}{m} + F\frac{\nabla P}{m}. \tag{22}$$

In reality, this behavior is not entirely unexpected, since, as we saw earlier, the actual velocity is the end result of the system dynamics as a whole, that is, S is also dictated by the local activities. Therefore, to reconcile these equations, F shall be such that (9) refers to the real part of (22).

The resulting apparent complex character of the energy (11) and the momentum (12) is only a stage of the calculations. In effect, the canonical equations (16) and (17) can also be obtained even making

$$\int d^3rP\left(-F\frac{\partial P}{\partial t}\right) = 0 \tag{23}$$

in Eq. (14), which makes the average energy (14) real. However, this implies that, on average, the exchange of energy between the particle and the field is zero, meaning that the energy provided by field is promptly returned to it in equal amount. This, besides constituting the desired local energy balance — it can be related with atomic stability [9] — also puts some insight in the complex shape of the mentioned real quantities.

In fact, the local energy balance (23) is satisfied by the trivial solution of (20), expressed by

$$F = \frac{\partial S_l}{\partial P} = \imath\frac{S_o}{P}, \tag{24}$$

as can be easily verified from the normalization of P. So this proven the P^{-1} functionality, which is not achieved only from probability conservation, as pointed above.

Substituting (24) into (19), results in

$$\bar{H} = \int d^3rP\left(\frac{(\nabla S)^2}{2m} + \frac{S_o^2}{2m}\frac{(\nabla P)^2}{P^2} + V\right), \tag{25}$$

which, with the canonical equations (16) and (17), reproduces the equations (3) and (4), respectively, if S_0 is identified with

$\hbar/2$. Therefore, to complete the classical derivation of the Schrödinger equation, the anzatz (2) must also be obtained in a classical context. This is the subject of the next section.

4 Parameterization of the equations

Knowing any one of the solutions (one of the paths) of the motion equations resulting from (25), the energy and momentum at each position, according to the equations (13), (12) and (24), are, respectively, given by

$$H = -\frac{\partial S}{\partial t} - \frac{\imath S_o}{P}\frac{\partial P}{\partial t} \tag{26}$$

and

$$\mathbf{p} = \nabla S + \frac{\imath S_o}{P}\nabla P. \tag{27}$$

Integrating these partial differential equations (minus a possible constant), we obtain the following dimensionless equation

$$\frac{1}{2\imath S_o}\left(\sum_i \int_0^{x_i} p_i dx_i - \int_0^t H dt\right) = \frac{S}{2\imath S_o} + \ln\sqrt{P}, \tag{28}$$

as can be easily verified by following the inverse procedure. The upper limits of the integrals are the coordinates and time of the positions occupied by the particle along a possible path, therefore the left hand side of (28) is a complex function of these parameters, which will be defined in the following way:

$$\ln\psi = \frac{1}{2\imath S_o}\left(\sum_i \int_0^{q_i} p_i dq_i - \int_0^t E dt\right). \tag{29}$$

As both sides of (28) are independent of the path followed by the particle, we can write the following relation between S and P, valid for all paths:

$$\ln\psi = \frac{S}{2\imath S_o} + \ln\sqrt{P}, \tag{30}$$

or

$$\psi = \sqrt{P}\exp\left(\frac{S}{2\imath S_o}\right). \tag{31}$$

This equation with $S_0 = \hbar/2$ is in full agreement with (2). And more, for constant energy and momentum the function defined in (29) is a solution of the Schrödinger equation for a free particle.

Finally, let's re-write the equations obtained in this work in terms of ψ. From (30) and its complex conjugate we obtain the following parametric shapes for S and P:

$$S = \frac{\imath\hbar}{2}\left(\lg\psi - \ln\psi^*\right) \tag{32}$$

and

$$P = \psi^*\psi. \tag{33}$$

Consequently, the equations (25), (26) and (27) can be re-written, respectively, in the forms:

$$\bar{H} = \int d^3 r\left(\frac{\hbar^2}{2m}\nabla\psi^*\cdot\nabla\psi + \psi^* V\psi\right), \tag{34}$$

$$\imath\hbar\frac{\partial\psi}{\partial t} = H\psi, \tag{35}$$

and

$$-\imath\hbar\nabla\psi = \mathbf{p}\psi. \tag{36}$$

Applying the divergence operator on both sides of equation (36), allied to fact that $\mathbf{p}$ is coordinate independent (it is independent of the followed path), gives

$$-\imath\hbar\nabla\cdot\nabla\psi = \mathbf{p}\cdot\nabla\psi, \tag{37}$$

and expressing $\mathbf{p}$ in terms of the complex conjugate of (36), we obtain the equality

$$-\psi^*\nabla^2\psi = \nabla\psi^*\cdot\nabla\psi. \tag{38}$$

Therefore the equation (34) can be written in the well known quantum form

$$\bar{H} = \int d^3 r\psi^*\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi. \tag{39}$$

5 Conclusion

The approach shows that the Schrödinger equation and its accessory are necessary and natural equations, parameterized shapes of the complicated — not to say unsolvable — equations resulting from a classical treatment including a special field with homogeneous, isotropic and random characteristics.

Submitted on May 21, 2011 / Accepted on May 28, 2011

References

1. Bohm D. A suggested interpretation of the quantum theory in terms of "hidden variables" I. *Physical Review*, 1952, v. 85, no. (2), 166–179.
2. Nerukh D., Frederick J. H. Multidimensional quantum dynamics with trajectories: a novel numerical implementation of the Bohmian mechanics. *Chemical Physics Letters*, 2000, v. 332, 145–153.
3. Holland P. V. The Quantum Theory of Motion. Cambridge University. Press, Cambridge, 1993.
4. Feynman R. P., Hibbs A. R. Quantum Mechanics and Path Integrals. McGraw-Hill Book Company 1965.
5. Milonni P. W. The Quantum Vacuum: An Introduction to Quantum Electrodynamics. Academic Press, 1994.
6. de la Peña L., Cetto A. M. The Quantum Dice. An Introduction to the Stochastic Electrodynamics. Kluwer Academic, Dordrecht, 1996.
7. Casimir H. B. G. On the attraction between two perfectly conducting plates. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, 1948, v. B51 793–795.
8. Lamoraux S. K. Demonstration of the Casimir Force in the 0.6 to 6μm Range. *Physical Review Letters*, 1997, v. 78, no. 1, 5–8.
9. Puthoff H. E. Ground state of hydrogen as a zero-point-fluctuation-determined state. *Physical Review D*, 1987, v. 35, no. 10, 3266–3269.

An Analysis of States in the Phase Space: the Anharmonic Oscillator

Sebastiano Tosto

ENEA Casaccia, via Anguillarese 301, 00123 Roma, Italy. E-mail: sebastiano.tosto@enea.it, stosto@inwind.it

The paper introduces a simple quantum model to calculate in a general way allowed frequencies and energy levels of the anharmonic oscillator. The theoretical basis of the approach has been introduced in two early papers aimed to infer the properties of quantum systems exploiting the uncertainty principle only. Although for clarity the anharmonic oscillator is described having in mind the lattice oscillations of atoms/ions, the quantum formalism of the model and approach have general character and can be extended to any oscillating system. The results show that the harmonic energy levels split into a complex system of anharmonic energy levels dependent upon the number of terms of the Hamiltonian that describes the anharmonicity.

1 Introduction

The anharmonic phenomena, well known in physics [1], regard a wide range of properties of practical and theoretical interest; e.g. in acoustics they account for large variations of sound velocity in solids [2], in optics for non-linear interaction of powerful light with lattice vibrations [3]. Moreover are known physical effects that lead to a behavior impossible in harmonic oscillators, like the "foldover effect" [4] and "superharmonic resonance" [5]; both are due to the dependency of the eigenfrequency of nonlinear oscillators on the amplitude and to the non-harmoniticity of the oscillations. In solid state physics, non-linear effects occur when atoms consisting of a positively charged nucleus surrounded by a cloud of electrons are subjected to an electric field; the displacement of nucleus and electrons causes an electric dipole moment, whose interaction with the applied field is linear for small field intensities only [6].

The present paper aims to propose a quantum mechanical approach to tackle the problem of non-harmonic oscillations in a general way, i.e. regardless of the particular issue of specific interest, and in line with the concepts introduced in two papers [7, 8] concerning simple quantum systems, many-electron atoms/ions and diatomic molecules. The basic idea of these papers starts from a critical review of the concepts of positions and momenta of interacting particles in a quantum system, where the dynamical variables are perturbed in a complex way by mutual interactions and change within appropriate ranges of values in agreement with boundary conditions like the minimum total energy.

Consider for instance the hydrogenlike atoms. It is reasonable to regard radial momentum p_ρ and distance ρ between electron and nucleus as variables included within proper ranges of values; so it is certainly possible to write $0 < \rho \leq \Delta\rho$ and $0 < p_\rho \leq \Delta p_\rho$ if $\Delta\rho$ and Δp_ρ have arbitrary sizes, including even the chance of infinite sizes. The only basic hypothesis of the quoted papers was that in general any ranges of conjugate dynamical variables Δx and Δp_x have physical meaning of quantum uncertainty ranges, thus to be regarded according to the basic ideas of quantum statistics; hence

$$\Delta x \Delta p_x = n\hbar, \tag{1.1}$$

with n arbitrary integer.

No hypothesis is necessary about Δx and Δp_x, which are by definition arbitrary, unknown and unpredictable. Eq. (1.1) was the unique assumption in [7, 8] and does so also in the model proposed here. Despite the apparently agnostic character of eq. (1.1), the results inferred in the quoted papers were in all cases completely analogous to that of the usual wave mechanics formalism; in particular it was found that the quantum numbers actually coincide with the numbers of allowed states in the phase space for the concerned systems. Eq. (1.1) only is enough to give the classical Hamiltonian, H_{cl}, the physical meaning of quantum Hamiltonian, H_q; it simply requires considering the ranges of dynamical variables rather than the dynamical variables themselves, which are therefore disregarded since the beginning. For instance, in a one-dimensional problem like that of a mass constrained to oscillate along a fixed direction, it means that hold the positions

$$H_{cl}(x, p_x) \Rightarrow H_q(\Delta x, \Delta p_x) \Rightarrow H_q(\Delta p_x, n). \tag{1.2}$$

The uncertainty is regarded in this way as fundamental principle of nature rather than as mere consequence of commutation rules of quantum operators. The case of the harmonic oscillator, already introduced in [7], has central importance here; its quantum formulation according to eq. (1.1) and positions (1.2) is so short and simple that it is sketched in the next section 2 to make the present paper clearer and self-consistent.

The next section aims also to emphasize how the concepts so far introduced enable the quantum approach. For clarity the anharmonic oscillator is regarded in section 3 having in mind the lattice oscillations of atoms/ions, yet through a very general approach that can be extended to any quantum system. The discussion on the results of the model and the conclusion are reported in sections 4 and 5.

2 The harmonic oscillator

With the positions 1,2, the classical energy equation $E = p^2/2m + k_{har}(x - x_o)^2/2$ of the oscillating mass around the equilibrium position x_o reads $\Delta E = \Delta p^2/2m + k_{har}\Delta x^2/2$, having omitted for simplicity the subscript x; owing to eq. (1.1), $E = E(\Delta p, n)$ is now because of n a random quantity within an energy range ΔE that corresponds to local uncertainty of dynamical variables within Δx and Δp. Both these latter and ΔE are assumed positive by definition. Then, one finds

$$\Delta E = \frac{\Delta p^2}{2m} + \frac{m(n\hbar\omega_{har})^2/2}{\Delta p^2}, \qquad \omega_{har}^2 = \frac{k_{har}}{m}. \tag{2.1}$$

Eq. (2.1) has a minimum as a function of Δp, i.e.

$$\Delta p_{\min} = \sqrt{mn\hbar\omega_{har}}, \qquad \Delta E_{\min} = n\hbar\omega_{har}, \tag{2.2}$$

being now n the number of vibrational states. Although for $n = 0$ there are no vibrational states, the necessity that $\Delta p \neq 0$ compels $\Delta E \neq 0$ and thus $\Delta E_0 = \Delta p_0^2/2m \neq 0$ with $\Delta p_0 = \Delta p_{\min}(n = 0)$. In this particular case, the problem reduces to that of a free particle in the box, i.e. Δp_0 is related to the zero point energy. This requires $\Delta p_0 = \Delta p_{\min}(n = 1)$, because the minimum quantum uncertainty of Δp can be nothing else but that of $\Delta p_{\min}$ for $n = 1$. The numerical correspondence between non-vibrational momentum range, Δp_0, and first vibrational momentum range, $\Delta p_{\min}(n = 1)$, means that at the zero point energy state the mass m is delocalized in a space range, $\Delta x_0 = \Delta x(n = 0)$, equal to that, $\Delta x(n = 1)$, pertinent to the lowest vibrational state. In other words, the oscillation amplitude at the ground energy level is the same as the delocalization range size of the particle with zero point energy only. Hence $\Delta p_0 = \sqrt{m\hbar\omega_{har}}$ defines $E_0 = \Delta p_0^2/2m = \hbar\omega/2$. The minimum of ΔE must be $\Delta E_{\min} = E_{\min} - \hbar\omega_{har}/2$; then, regarding $E_{\min} = E_{har}$ as the harmonic energy level, the known result

$$E_{har} = n\hbar\omega_{har} + \frac{\hbar\omega_{har}}{2} \tag{2.3}$$

is obtained considering uncertainty ranges of eq. (1.1) only, and without any further hypothesis. Note that with $\Delta p = \Delta p_{har}$

$$\frac{\Delta p^2}{2m} = \frac{\omega_{har}^2 mn^2\hbar^2}{2\Delta p^2} = \frac{n\hbar\omega_{har}}{2},$$

in agreement with the virial theorem as $E_{\min}$ is given by the sum of kinetic and potential terms, whereas the zero point term has kinetic character only. Also note in this respect that $\Delta p_{\min}$ and Δp_0 are merely particular range sizes, among all the ones allowed in principle, fulfilling the condition of minimum $E_{\min}$ and E_0.

These results do not contradict the complete arbitrariness of Δp and Δx, since in principle there is no compelling reason to regard the particular ranges of eqs. (2.2) in a different way with respect to all the other ones allowed by eq. (1.1); rather the results merely show the preferential propensity of nature for the states of minimum energy. In effect it is not surprising that the energy calculated with extremal values of dynamical variables in the ranges of eq. (2.1) does not coincide, in general, with the most probable energy. In conclusion, this example highlights that the physical properties of a quantum system can be inferred without solving any wave equation simply replacing the local dynamical variables with the respective quantum uncertainty ranges: the key problem becomes then that of counting correctly case by case the appropriate number of allowed states, as shown in [7,8] for more complex quantum systems.

It appears that, once accepting the eq. (1.1) and the consequent positions 1,2, have actual physical meaning the uncertainty ranges rather than the dynamical variables themselves; these latter are considered here random, unknown and unpredictable within the respective ranges and thus are disregarded since the beginning when formulating the physical problem. Just this is the essence of eq. (2.1). Eventually note that the vibrational quantum number n appears to be here the number of quantum states allowed to the oscillator. Since the present approach gives sensible results for harmonic oscillations, there is no reason to exclude that the same holds for anharmonic oscillations as well. The next paragraph aims to generalize the kind of approach just introduced to the case of anharmonic oscillations.

3 The anharmonic oscillator

The classical Hamiltonian reads now

$$E = p^2/2m + \sum_{i=2}^{N} a_i'(n\hbar)^{-i}(x - x_o)^i, \tag{3.1}$$

being N the arbitrary number of terms of the series including quadratic and anharmonic terms and a_i' proper coefficients assumed known; indeed the values of these coefficients characterize distinctively the specific kind of oscillating system. The signs of a_3' and a_4' are taken here negative [9]; the former expresses the asymmetry of the mutual repulsion between atoms or ions, e.g. in a metallic lattice, the latter describes the softening of the vibration at large amplitudes. The higher order terms allow to describe these effects in a more general way, so their sign and values must agree with the idea that the global consequence of anharmonicity is to lower the potential energy of oscillation; indeed the potential energy reads $a_2'(x - x_o)^2 f(x)$, i.e. it consists of a quadratic term with x-dependent correction factor $f(x) = 1 + \sum_{i=3}^{N}(a_i'/a_2')(x - x_o)^{i-2} < 1$. By analogy with the harmonic case, the coefficient of the quadratic term, anyway related to the force constant k_{an}, is reasonably expected to have still the form $m(n\hbar\omega_{an})^2/2$ with oscillation frequency defined now by $\omega_{an}^2 = k_{an}/m$. Moreover the dependence of this term on ω_{an} suggests that in general $a_i' = a_i'(\omega_{an})$ are to be expected as well.

The following discussion aims to guess this dependence and the relationship between ω_{an} and ω_{har} through the same

approach shown previously; so, as done in section 2, we aim to calculate $\Delta E_{\min}$ and infer next the anharmonic vibrational levels E_{an} and zero point energy E_0, being clearly $\Delta E_{\min} = E_{\min} - E_0$ and $E_{an} = E_{\min}$.

According to the position (1.2) and eq. (1.1), the quantum energy equation corresponding to eq. (3.1) reads

$$\Delta E = \frac{\Delta p^2}{2m} + \sum_{i=2}^{N} \frac{a'_i}{\Delta p^i}. \tag{3.2}$$

This equation, minimized with respect to the range Δp, yields

$$\Delta p_{\min} = m \sum_{i=2}^{N} i a'_i \Delta p_{\min}^{-(i+1)}, \tag{3.3}$$

being

$$\Delta E_{\min} = \Delta E(\Delta p_{\min}), \qquad \Delta p_{\min} = \Delta p_{\min}(\omega_{an}).$$

For assigned coefficients a'_i, the first equation admits in general $N + 2$ solutions $\Delta p_{\min}$, some of which can be however imaginary. Being the momentum uncertainty range Δp real positive by definition, let $I' \leq N + 2$ be the number of positive real roots; so I' possible values of $\Delta p_{\min}$ describe the allowed momentum ranges of the oscillating particle that fulfil the minimum condition. A further limitation to these values is that the series must converge. Disregard also the values of $\Delta p_{\min}$ that with the given a'_i possibly do not fulfil the inequality $\left|(i+1)a'_{i+1}\Delta p_{\min}^{-(i+2)}\right| << \left|ia'_i\Delta p_{\min}^{-(i+1)}\right|$ inferred from eq. (3.2), i.e.

$$\left|a'_{i+1}\right| << \left|a'_i \Delta p_{\min}\right|. \tag{3.4}$$

Then $I \leq I'$ is the number of real roots of physical interest to be considered in the following. Trivial manipulations of eq. (3.2) to eliminate m with the help of eq. (3.3) yield

$$\Delta E = \frac{1}{2}\left(\frac{\Delta p}{\Delta p_{\min}}\right)^2 \sum_{i=2}^{N} \frac{i a'_i}{\Delta p_{\min}^i} + \sum_{i=2}^{N} \frac{a'_i}{\Delta p^i}. \tag{3.5}$$

To extract the allowed physical information from this equation one should minimize with respect to Δp and then proceed as shown in the harmonic case. Actually this minimum condition has been already exploited to infer eq. (3.3) itself, which suggests that eq. (3.5) should not need being minimized once more. To understand this point replace Δp with $\Delta p_{\min}$ in eq. (3.5) and consider first the resulting equation $\Delta E(\Delta p_{\min}) = \sum_{i=2}^{N}(1 + i/2)a'_i\Delta p_{\min}^{-i}$ in the harmonic case; then $N = 2$, i.e. $a'_{i>2} = 0$, yields $3a'_2\Delta p_{\min}^{-2}/2$. By comparison with eq. (2.1) this result takes a more familiar form replacing a'_2 with $a_2\Delta p_{\min}^4/m$ where a_2 is a dimensionless proportionality coefficient linking a'_2 and $\Delta p_{\min}$; in this way one obtains

$$\Delta E(\Delta p_{\min}) = \frac{3a_2\Delta p_{\min}^2}{2m},$$

which has the same form of eqs. (2.2) a proportionality factor apart. As expected, an immediate connection with the harmonic case is possible uniquely on the basis of the condition 3,3 without introducing explicitly neither ω_{har} nor the equations of Δp_{har} and ΔE_{har}. Express thus in general the coefficients a'_i as a function of $\Delta p_{\min}$ as follows

$$a'_i = \frac{\Delta p_{\min}^{i+2}}{m} a_i, \qquad \sum_{i=2}^{N} i a_i = 1, \qquad 1 \leq j \leq I \tag{3.6}$$

where a_i are new constants that fulfil the boundary condition expressed by the second equation, straightforward consequence of eq. (3.3). Note that a'_i are uniquely defined for the specific oscillating system, whereas the appropriate notation of the various a_i should be $a_i^{(j)}$ to emphasize that a set of these coefficients is defined by each solution $\Delta p_{\min}^{(j)}$ of physical interest calculated through eq. (3.3). This would also compel indicating in eq. (3.5) $\Delta E^{(j)}$ and then $\Delta E_{\min}^{(j)}$. To simplify the notations the superscript (j) will be omitted, stressing however once for all that if $N > 2$ then eq. (3.5) actually represents anyone among I admissible equations. Replacing a'_i into the energy equation (3.5), one finds

$$\Delta E = \left(\left(\frac{\Delta p}{\Delta p_m}\right)^2 \frac{1}{2} + \sum_{i=2}^{N} a_i \left(\frac{\Delta p_m}{\Delta p}\right)^i\right) \frac{\Delta p_{\min}^2}{m}.$$

This suggests putting

$$q\frac{\Delta E}{\Delta E_{\min}} = \frac{1}{2}\left(\frac{\Delta p}{\Delta p_{\min}}\right)^2 + \sum_{i=2}^{N} a_i \left(\frac{\Delta p_{\min}}{\Delta p}\right)^i, \tag{3.7}$$

$$\Delta E_{\min} = q\frac{\Delta p_{\min}^2}{m}, \qquad a_2 = \frac{1}{2}\left(1 - \sum_{i=3}^{N} i a_i\right).$$

The proportionality factor q aims to fulfil the reasonable condition $\Delta E = \Delta E_{\min}$ for $\Delta p = \Delta p_{\min}$ and express in a general way the expected link between $\Delta E_{\min}$ and $\Delta p_{\min}^2/m$. Trivial calculations yield

$$q = 1 + \sum_{i=3}^{N}(1 - i/2)a_i. \tag{3.8}$$

Of course q must be intended here as $q^{(j)}$ likewise as $a_i^{(j)}$. Whatever a_i might be, eq. (3.7) does not need being minimized; it simply expresses as a function of $\Delta p/\Delta p_{\min}$ the energy deviation from the harmonic condition for assigned values of the coefficients $a'_{i\geq 2} \neq 0$. Eq. (3.7) and a_2 are uniquely defined in the particular case $a_{i>2} = 0$ only, which corresponds to $q = 1$ as well. Moreover the form of the second equation, analogous to that of eqs. (2.2), suggests that $\Delta p_{\min}$ and $\Delta E_{\min}$ must be also equal or proportional to the respective harmonic quantities Δp_{har} and ΔE_{har}. So putting

in general $\Delta E_{\min} = w^2 \Delta E_{har}$ and $\Delta p_{\min} = w \Delta p_{har}$, with w proportionality factor, one finds

$$\frac{q}{w^2}\frac{\Delta E}{\Delta E_{har}} = \frac{1}{2w^2}\left(\frac{\Delta p}{\Delta p_{har}}\right)^2 + \sum_{i=2}^{N} a_i w^i \left(\frac{\Delta p_{har}}{\Delta p}\right)^i, \qquad (3.9)$$

$$\omega_{an} = w^2 \omega_{har}.$$

Likewise q, also w must be intended in general as $w^{(j)}$. So eqs. (3.9) define I anharmonic frequencies $\omega_{an}^{(j)} \neq \omega_{har}$, here designated shortly ω_{an}, corresponding to the unique harmonic frequency ω_{har}; i.e. the various $\Delta E_{\min}$ describe the splitting of each n-th vibrational energy level $n\hbar\omega_{har}$. The anharmonic potential of eq. (3.9) is expected to depend upon ω_{an} through the dimensionless coefficients a_i, by analogy with the dependence of the harmonic term upon ω_{har}^2. Thus, to complete the task of the present section it is necessary: (i) to define the factor w of eq. (3.9); (ii) to highlight the analytical form of the functions $a_i = a_i(\omega_{an})$; (iii) to express the potential energy of equation (3.9) as a function of ω_{an} through these coefficients.

Rewrite to this purpose the coefficients of eq. (3.2) as shown in following series

$$q\Delta E = \frac{1}{2}\frac{\Delta p^2}{m} + \sum_{i=2}^{N} a_i'' \frac{m^{i/2}(n\hbar\omega_{an})^{i/2+1}}{\Delta p^i}, \qquad (3.10)$$

where the powers of $n\hbar\omega_{an}$ and m have been determined by dimensional consistency of the various terms with both ΔE and Δp^i. Minimizing with respect to Δp and equating to zero, one finds

$$R_E = \frac{1}{2}R_p^2 + \sum_{i=2}^{N} a_i'' R_p^{-i}, \qquad (3.11)$$

where

$$R_E = q\frac{\Delta E}{\Delta E_{har}}, \qquad R_p = \frac{\Delta p}{\Delta p_{har}}, \qquad a_i'' = a_i w^{i+2}.$$

With the coefficients a_i'' and a_i linked by the last position, eq. 3,(11) is identical to eq. (3.9); this consistency supports therefore the positions of both eqs. (3.6) and (3.10). To specify w put first $N = 2$ in eq. (3.9); minimizing $R_p^2/2w^2 + a_2 w^2/R_p^2$ with respect to R_p yields $R_p^4 = 2a_2 w^4$. Since the minimum of R_p can be nothing else but 1 by definition, $w = (2a_2)^{-1/4}$ yields $w = 1$, whereas in this particular case $a_2 = 1/2$. As expected, eq. (3.9) is thus uniquely defined for $a_{i>2} = 0$ only. Note that the coefficient of the quadratic term of eq. (3.10) reads $a_2'' m(n\hbar\omega_{an})^2$; if the result $w = (2a_2)^{-1/4}$ previously obtained for $N = 2$ still holds for any N with a_2 given now by the last eq. (3.7), then $a_2'' = a_2 w^4$ yields $a_2'' = 1/2$ and thus the expected form $m(n\hbar\omega_{an})^2/2$ formerly quoted whatever $a_{i>2}$ might be.

This consideration encourages one to conclude with the help of eq. (3.7)

$$w^2 = (2a_2)^{-1/2} = \left(1 - \sum_{i=3}^{N} i a_i\right)^{-1/2},$$

$$a_i'' = a_i \left(1 - \sum_{i=3}^{N} i a_i\right)^{-i/4-1/2}.$$

Replacing a_i'' in eq. (3.10) one finds

$$\Delta E = \frac{1}{2q}\frac{\Delta p^2}{m} + \\ + \sum_{i=2}^{N} q^{-1} a_i \frac{m^{i/2}(n\hbar\omega_{har})^{i/2+1}}{\Delta p^i}\left(1 - \sum_{i=3}^{N} i a_i\right)^{-\frac{3}{4}(i+2)}. \qquad (3.12)$$

This is the sought generalization of eq. (2.1) when $a'_{i>2} \neq 0$; the positions so far introduced link eq. (3.2) with the harmonic case. Moreover eq. (3.9) yields

$$\omega_{an} = \left(1 - \sum_{i=3}^{N} i a_i\right)^{-1/2} \omega_{har}. \qquad (3.13)$$

With the given choice of w^2, therefore, $a_{i\geq 3} = 0$ yield not only $\omega_{an} = \omega_{har}$ but also $\Delta p_{\min} = \Delta p_{har}$ and $\Delta E_{\min} = \Delta E_{har}$. Hence

$$\Delta E_{\min} = n\hbar\omega_{an} = \left(1 - \sum_{i=3}^{N} i a_i\right)^{-1/2} n\hbar\omega_{har}, \qquad (3.14)$$

$$\Delta p_{\min} = \sqrt{mn\hbar\omega_{an}} = \left(1 - \sum_{i=3}^{N} i a_i\right)^{-1/4} \sqrt{mn\hbar\omega_{har}}.$$

As concerns the zero point energy E_0 hold the considerations of the previous section, i.e. $\Delta E_{\min} = E_{\min} - E_0$; moreover also now for $n = 0$ the minimum of eq. (3.12) reduces to $\Delta p_0^2/2qm$, with $\Delta p_0^2 = \Delta p_{\min}^2 (n = 0)$. As explained before, even in lack of vibrational states $\Delta p_{\min} \neq 0$ compels putting $\Delta p_0 = \Delta p_{\min}^{(0)} (n = 1)$ by virtue of eq. (3.14) so that $E_0 = \left(1 - \sum_{i=3}^{N} i a_i^{(0)}\right)^{-1/2} \hbar\omega_{har}/2q$; since in general are allowed several values of $\Delta p_{\min}$, the notation emphasizes that one must consider here the set of values of $a_i^{(j)}$ corresponding to the smallest among the various $\Delta p_{\min}^{(j)}$.

In conclusion, since the anharmonic energy and momentum must correspond to the respective $\Delta E_{\min}$ and $\Delta p_{\min}$, it is possible to summarize the previous results, with full notation for clarity, as follows with the help of eq. (3.8)

$$E_{an}^{(j)} = \left(1 - \sum_{i=3}^{N} i a_i^{(j)}\right)^{-1/2} n^{(j)}\hbar\omega_{har} + \\ + \frac{1}{2}\left(1 + \sum_{i=3}^{N}(1 - i/2) a_i^{(0)}\right)^{-1} \left(1 - \sum_{i=3}^{N} i a_i^{(0)}\right)^{-1/2} \hbar\omega_{har}, \qquad (3.15)$$

with

$$1 \leq j \leq I,$$

$$\Delta p_{an}^{(j)} = \left(1 - \sum_{i=3}^{N} i a_i^{(j)}\right)^{-1/4} \sqrt{m n^{(j)} \hbar \omega_{har}},$$

$$\omega_{an}^{(j)} = \left(1 - \sum_{i=3}^{N} i a_i^{(j)}\right)^{-1/2} \omega_{har},$$

$$a_i^{(j)} = \frac{m a_i'}{\left(\Delta p_{\min}^{(j)}\right)^{i+2}}.$$

4 Discussion

The strategy of the papers [7, 8] to exploit via eq. (1.1) the classical Hamiltonians of the system of interest was outlined in section 2 and then extended in section 3 to the anharmonic case. The first task of the discussion aims to clarify the classical and quantum ways to regard the harmonic and anharmonic oscillation. The classical potential energy of eq. (3.1), $U_{cl} = U_{cl}(x - x_o)$, concerns a withholding force progressively increasing as a function of $x - x_o$ while the oscillation turns gradually from harmonic into anharmonic behaviour. Moreover if momentum and position of m are both exactly known, U_{cl} can be defined with arbitrary accuracy simply increasing the number of terms of the series.

This description is clearly inadequate for the potential energy, $U_q = U_q(\Delta p, n)$, of the quantum eq. (3.2); in principle the exact elongation of m with respect to the rest position and the corresponding momentum are not jointly specifiable, i.e. the limit $\Delta x \to 0$ could not be described by finite values of $\Delta p_{\min}$. Indeed $\Delta p_{\min} \to \infty$ compels $\Delta p \to \infty$ that yields $\Delta E = \Delta p^2/2m$ regardless of a_i'; this limit corresponds to the classical case of a free particle in a one-dimensional box, of no interest here, rather than to the harmonic limit expected for $\Delta x \to 0$. Eventually the quantum uncertainty compels regarding in a different way also the number of terms of U_{cl} and of U_q: in the former case N is in principle arbitrary, being significant its actual ability to provide a description as detailed as possible of the local state of motion of m, in the latter case does not, being instead significant its actual ability to introduce the allowed physical information into the system.

If for instance the model aims to describe softening and asymmetry effects only, then are justified terms like Δx^i with powers and signs [9] pertinent to these effects only. Solving eq. (3.1) requires exploiting the functional relationship U_{cl} upon Δx through numerical methods, solving eq. (3.2) requires instead a different reasoning because the anharmonic effects inherent the various Δx^i are related to the respective Δp^{-i} through eq. (1.1) only: the previous results show that a general physical principle, the minimum energy, is enough to this purpose. According to the classical eq. (3.1) the harmonicity requires $a'_{i\geq 3}\Delta x^i << a_2'\Delta x^2$ in agreement with the convergence condition (3.4); the quantum eq. (3.2) requires $a'_{i\geq 3}\Delta p^{-i} << a_2'\Delta p^{-2}$, which is still a statement of "small" oscillation amplitudes since $a_i'\Delta p^{-i} \propto a_i'\Delta x^i$. Both definitions are thus equivalent, yet the latter is more interesting because it involves eq. (1.1) and allows further considerations on the classical and quantum concepts of harmonicity. Eq. (3.4) and the first eq. (3.3) yield for $i \geq 3$

$$a'_{i\geq 3}\Delta p^{-i} << a_2'\Delta p^{-2} \Rightarrow a_i \left(\frac{\Delta p_{\min}}{\Delta p}\right)^i << a_2 \left(\frac{\Delta p_{\min}}{\Delta p}\right)^2.$$

Noting that Δp is arbitrary by assumption and that $\Delta p_{\min} \leq \Delta p$ by definition, it turns out that the second inequality can be merely fulfilled by $\Delta p/\Delta p_{\min} >> 1$ regardless of the values of the ratios a_2/a_i and a_2'/a_i'. Since in principle a_i' only are required to fulfil the convergence condition (3.4), whereas the values of a_i are ineffective in this respect because their values are consequently defined in the successive eq. (3.6) only, the conclusion is that small oscillation amplitudes do not require necessarily vanishing $a_{i>2}$. According to eq. (3.13), however, just these latter define w that in turn control ω_{an} and thus the splitting of energy levels. The fact that in general $w \equiv w^{(j)} \neq 1$ even for small oscillations supports the idea that the quantum harmonicity is a particular case, but not a limit case, of the quantum anharmonicity; in other words, an oscillating quantum system does not change gradually from harmonic to anharmonic behaviour.

This conclusion is confirmed also considering the dependence of the constants w on a_i. In eq. (3.6) large values of $\Delta p_{\min}$ entail small a_i and thus w such that the corresponding allowed frequencies ω_{an} are expected to have values similar to ω_{har}; the contrary holds for small values of $\Delta p_{\min}$, to which correspond larger values of w and therefore larger gaps $\omega_{an} - \omega_{har}$.

Hence, when considering the totality of allowed frequencies consistent with the different sizes of all ranges $\Delta p_{\min}$, even small values of a_i' classically compatible with the harmonic condition entail anyway relevant splitting and gap of energy levels with respect to ω_{har} typical of the anharmonicity; otherwise stated, the quantum harmonicity requires $a'_{i\geq 3} = 0$ exactly. The harmonic ground level is a reference energy rather than an attainable limit energy because fails the classical expectation of anharmonic frequencies progressively deviating from ω_{har} along with a_i'; the last eq. (3.7) shows indeed that even the first quadratic coefficient a_2 of potential energy differs from the corresponding harmonic coefficient unless $a_{i\geq 3} = 0$. It is also significant the fact that the unique ω_{har}, classically defined in eq. (2.1) through the force constant k_{har} of Hooke law only, never corresponds to a unique ω_{an} whatever $a'_{i\geq 3} \neq 0$ might be; this latter, although formally introduced in the early eq. (3.3) as $\omega_{an}^2 = k_{an}/m$, has quantum character after being subsequently redefined by eq. (3.9) through the multiplicity of values of w.

It is however worth noting in this respect a further chance to define the oscillation frequency in a mere quantum way through an uncertainty equation having a form seemingly different but conceptually equivalent to eq. (1.1). Introduce the

time range Δt necessary to displace m by Δx with finite average velocity v_x; defining then $\Delta t = \Delta x/v_x$ and $\Delta E = \Delta p v_x$, eq. (1.1) takes a form that introduces new dynamical variables t and E having random, unpredictable and unknown values within the respective uncertainty ranges defined by the same $n\hbar$. Of course Δt and ΔE are completely arbitrary, as they must be, likewise Δx and v_x. Thus, with the constrain of equal n, eq. (1.1) reads equivalently as

$$\Delta E \Delta t = n\hbar, \qquad \Delta t = t - t_o, \tag{4.1}$$

being the constant t_o the arbitrary origin of time coordinates. Eq. (4.1) is not a trivial copy of eq. (1.1): it introduces new information through v_x and shows that during successive time steps Δt the energy ranges ΔE change randomly and unpredictably depending on n. Of course the eq. (1.1) could have been inferred itself in the same way from eq. (4.1), i.e. regarding this latter as the fundamental statement. Relating eqs. (1.1) and (4.1) via the same arbitrary integer n, whatever it might be, means describing the oscillation of m through energy and time uncertainty ranges. This is equivalent to say that the time coordinate is regarded in an analogous way as the space coordinate hitherto concerned.

To show the consequences of this assertion, consider that $1/\Delta t$ has in general physical dimensions of frequency; then eq. (4.1) can be rewritten as $\Delta E_n = n\hbar\omega^{\S}$, being $\omega^{\S}$ a function somehow related to any frequency ω. If in particular $\omega^{\S}$ is specified to be just the previous frequency ω_{har}, whatever the value of this latter might be, eq. (4.1) reads

$$\Delta E_n = n\hbar\omega_{har}. \tag{4.2}$$

The notation emphasizes that the particular case $\omega^{\S} \equiv \omega_{har}$ enables a direct conceptual link with eq. (2.3), i.e. it concerns the harmonicity; having found that n is according to eq. (1.1) the number of vibrational states of the oscillator and $n\hbar\omega_{har}$ their energy levels, then without need of minimizing anything one infers that ΔE_n is again the energy gap between the n-th excited state of the harmonic oscillator and its ground state of zero point energy; the condition of minimum energy and $\Delta p_{\min}$ are now replaced by the specific meaning of Δt.

This conclusion shows that a particular property of the oscillating system is correlated to a particular property of the uncertainty ranges, thus confirming the actual physical meaning of these latter. So E_n falling within ΔE_n are still now random, unpredictable and unknown because of n. While ω_{har} was formerly defined by the formal position 2,1, now eq. (4.1) reveals its actual quantum meaning due to its direct link with the time uncertainty Δt.

This last result is significant for the present discussion: it justifies the different outcomes of the quantum approach with respect to the classical expectation in terms of uncertainty about the dynamical variables of m only; thus, as shown in [7, 8], this result disregards any phenomenological/classical hint to describe the system. In other words, instead of thinking to a withholding spring bound to a mass moving back and forth, the oscillation can be imagined in a more abstract way. It is enough to introduce an arbitrary energy range ΔE_n to which corresponds a respective quantum frequency $1/\Delta t_n$; then the form of eq. (1.1) is suitable to introduce an appropriate potential energy with elongation extent described by a unique quadratic term or by a series of terms, whose coefficients are respectively expressed as a function of ω_{har} or ω_{an} like in eqs. (2.1) or (3.10).

The worth of this conclusion is due to the generality of the resulting concept of oscillation, which skips any information on actual kind of motion of m, particular property of the oscillating mass, specific nature of the withholding force and hypothesis on the allowed range of frequencies. Both time and space uncertainties allow thus to describe an oscillating system in a fully quantum way, without writing and solving its wave equation. The previous results highlight the link of the allowed frequencies to the terms of U_q, see in particular the remarks about eqs. (3.5) and 3,13. A consequence of this point of view is that replacing U_{cl} with U_q compels the existence of several momentum uncertainty ranges $\Delta p_{\min}$ and thus of as many ω_{an} even when one would expect a mere perturbation of the unique ω_{har}: the physical information provided by the quadratic term only is uniquely defined, instead the various values of $\Delta p_{\min}$ and ω_{an} for $N > 2$ in eq. (3.2) reveal according to the last eq. (3.7) multiple anharmonic effects that influence also the quadratic term. The quantum uncertainty is therefore crucial in describing the oscillation.

For instance let us show that, at least for certain frequencies, the anharmonic oscillator appears to be a system intrinsically unstable. Let i be the index of any high order term of the series such that $a'_i/\Delta p^i << a'_2/\Delta p^2$ is true by definition because of the convergence condition; so $a'_i/\Delta p^i$ represents a small contribution to the total energy of oscillation. Let $\delta a'_i/\Delta p^i$ be its value altered by the change of the coefficient a_i because of an external perturbation acting on the oscillator; if for instance an impurity diffuses through the lattice in proximity of the given oscillating atom/ion, the stress field around this impurity or its possible charge field reasonably modify the local repulsion between atoms/ions or the softening effects at large oscillation amplitudes, as a consequence of which the anharmonic coefficients a'_3 and/or a'_4 are expected to change.

Let us exemplify any perturbation like this through a suitable change of some a'_i of the i-th energy terms in eq. (3.2); here however we consider for simplicity one term only to describe the local effect. The proof that some $\Delta p_{\min}$ and resulting $\Delta E_{\min}$ are strongly affected even by a very small change of any $a'_{i>2}$ is easy in the particular case where the series describing the potential energy converges very quickly. Differentiating eq. (3.6) one finds

$$\delta a'_i = a'_i \left((i+2) \frac{\delta \Delta p_{\min}}{\Delta p_{\min}} + \frac{\delta a_i}{a_i} \right).$$

Fix the value of $\delta a_i'$; if the local perturbation of the lattice affects a_i' in such a way that $\delta a_i' >> a_i'$, i.e. it alters significantly a_i, then the quantity in parenthesis is very large. If this happens while holds for $\delta a_i'$ also the condition $\delta a_i'/\Delta p^i << a_2'/\Delta p^2$, still possible because no hypothesis has been made on the strength of the perturbation, then considering that the quadratic term provides the most essential contribution to the total potential energy the result is: even a small perturbation $\delta a_i'/\Delta p^i$ of the whole oscillation energy is able to change significantly both $\Delta p_{\min}$ and a_i that define ω_{an}, see eqs. (3.13) to (3.15). The altered size of the range $\Delta p_{\min}$, actually verified by preliminary numerical simulations carried out with coefficients a_i' arbitrarily chosen to match the aforesaid condition, means in particular that the whole energy of the system admits not only a different ω_{an} allowed to the oscillator but also a larger range of corresponding momenta $p_{\min}$ allowed to m; this does not exclude even the chance of chaotic motion related to a random sequence of values ω_{an} during a weak perturbation transient due to the diffusing impurity.

The reason of such instability rests once again on the different way of regarding the oscillation amplitudes in classical and quantum physics. The former admits the limit $\Delta x \to 0$ regardless of Δp, the latter does not; so the quantum oscillation range of physical interest cannot be arbitrarily small or change arbitrarily without violating the crucial condition of minimum energy. Indeed the oscillation range sizes corresponding to the vibrational levels are quantized themselves

$$\Delta x_{\min} = \sqrt{\frac{n\hbar}{\omega_{an} m}}, \qquad \Delta x_0 = \sqrt{\frac{\hbar}{\omega_{an}^{(0)} m}}.$$

At this point it is worth remembering what has been previously emphasized, i.e. that the sizes of the ranges Δx and Δp are unspecified and indefinable; $\Delta x_{\min}$ and $\Delta p_{\min}$ are merely particular values showing the propensity of nature to fulfil the condition of minimum energy, however without contradicting the assumption that the uncertainty ranges are in principle completely arbitrary. So oscillation ranges that do not fulfil the former condition are certainly possible but unstable because of mere quantum reasons, i.e. they do not correspond to momentum range sizes that minimize the oscillation energy levels.

This conclusion is important because its validity follows uniquely from the assumption of convergence of the potential series only, i.e. it concerns a realistic condition effectively possible for the oscillator rather than an unusual and improbable limit case. Also, this result holds whatever the origin of the anharmonicity might be and confirms the physical diversity of harmonic and anharmonic quantum systems. Note however that the former is actually an ideal abstraction only; what can be expected in practice is a strong or weak anharmonicity, unless some specific physical reason requires just a potential energy with quadratic term only. So the results of the present approach should be regarded as the realistic behaviour of any oscillating system, rather than a sophisticated improvement of the naive harmonic behaviour; now this latter appears thus in general reductive and incomplete, rather than merely approximate. Yet eq. (3.15) shows that the zero point energy is formally analogous in both cases, a numerical difference apart: the only difference between the harmonic and anharmonic cases is that instead of considering the unique $\hbar\omega_{har}/2$ one must select the smallest $\omega_{an}^{(j)}$ to calculate $\hbar\omega_{an}^{(0)}/2$.

Note eventually that easy considerations allow to generalize the concept of perturbed oscillator in the conceptual frame of the present model. So far the present approach aimed to introduce the terms a_3 and a_4 to account for the anharmonicity, so that eqs. (3.2) to (3.15) tacitly assume an isolated oscillating system. Simple considerations however allow to further generalize the physical meaning of eq. (3.2) taking advantage of the fact that the present model works with a number of high order terms in principle arbitrary. In particular coefficients and number of terms could be exploited to describe even an oscillating system perturbed by an external force, for instance due to the interaction with other oscillators; indeed this force can be certainly described as a series development having the form $\sum a_i'' \Delta x^i$ if it is related, in the most general case non-linearly, to the displacement extent of the oscillating mass. Of course i can be even negative if the force vanishes at infinite distance. So, whatever the nature of the perturbation might be, this means that the potential energy of the system changes by an additional amount $-\sum a_i''(1+i)^{-1}\Delta x^{i+1}$ to be summed up with the corresponding terms of eq. (3.1). In any case, however, adding an arbitrary number of such energy terms to those intrinsically characterizing the oscillator does not change in principle the approach so far exposed, except of course the numerical value of the various a_i of eq. (3.9), which are now replaced by the sum $a_i'' + a_i'$ for each i-th power of oscillation elongation. So nothing hinders to regard the energy range ΔE_{an} of this equation as $\Delta E_{an+pert}$ still normalized to that of an isolated harmonic oscillator; it is enough that the coefficients a_i' up to the N-th order are still known, i.e. defined by the particular kind of oscillating system and external perturbation, yet without necessarily assuming any constrain on their signs, now determined by the sum of both effects. Even in the case where the force is described by terms like $\alpha'/\Delta x^i$ one would find an equation like (3.2) containing however terms like $a_k'\Delta p^k$ with $k > 2$. Also in this case, minimizing with respect to Δp would yield an appropriate number of roots $\Delta p_{\min}$ and thus prospective conclusions in principle completely analogous to that previously carried out. In the present case holds therefore the following position

$$\omega_{an+pert} \gtreqless w^2 \omega_{har}.$$

As expected, the previous scheme of vibrational levels is modified the external perturbation that affects w. This last result confirms the very general character of the way to describe any oscillating system simply with the help of the fundamental eq. (1.1).

5 Conclusion

The computational scheme introduced in the present paper is very simple: the most important achievements hitherto exposed do not require numerical calculations, but are consequence of general considerations on basic concepts of quantum mechanics. The general character of the approach, e.g. due to the arbitrary number N of anharmonic terms, and the possibility of extension to the case of a perturbed oscillator, propose the model as a useful tool in a broad variety of physical problems.

Submitted on June 13, 2011 / Accepted on June 16, 2011

References

1. Landau L. D., Lifshitz E. M., Mecanique, (1969), Editions MIR, Moscou, p. 118 and ff.
2. Hoechli U. T. Acoustic hanarmonicity and the ΔE effect in tetragonal $SrTiO_3$. *Solid State Communications*, 1973, v. 13, no. 9, 1369–1373.
3. Piekara A. H., Ratajska B. Nonlinear interaction of a powerful light beam with crystal lattice vibrations. *Applied Optics*, 1978, v. 17, no. 5, 689–690.
4. Petrovsky A. B. Foldover ferromagnetic resonance and damping in permalloy microstrips. *Physics Letters A*, 1968, v. 27, 220–221.
5. Abdel-Rahman E. M., Nayfeh A. H. Superarmonic resonance of an electrically actuated resonant microsensor. *Proceedings of the International Conference on Nanotechnology, Nanotech*, 2003, v. 2, 440–443.
6. Berrondo M., Recamier J. Resonances and anti-bound states in a Morse potential. *Journal of Chemical Physics Letters*, 2011, v. 503, no. 1–3, 180–184.
7. Tosto S. An analysis of states in the phase space: the energy levels of quantum systems. *Il Nuovo Cimento B*, 1996, v. 111, no. 2, 193–215.
8. Tosto S. An analysis of states in the phase space: the diatomic molecules. *Il Nuovo Cimento D*, 1996, v. 18, no. 12, 1363–1394.
9. Kittel C. Introduction to Solid State Physics. Third Edition, J. Wiley and Sons, New York, 1967, p. 184.

Gravity and the Conservation of Energy

Douglas L. Weller
Email: physics@dougweller.com

The Schwarzschild metric apportions the energy equivalence of a mass into a time component, a space component and a gravitational component. This apportionment indicates there is a source of gravitational energy as well as a limit to the magnitude of gravitational energy.

1 Introduction

Albert Einstein asserted that his field equations are in essence a restatement of the conservation of energy and momentum [1, pp. 145–149]. Every solution of the field equations, therefore, must account for all energy in the system described by the solution. How do solutions to the field equations account for gravitational energy?

This paper explains how within Schwarzschild's solution [2] to Einstein's field equations the effects of gravity can be represented as a velocity and as an apportionment of mass-energy equivalence. This allows an accounting for gravitational energy as part of mass-energy equivalence.

The paper first considers a spacetime without gravity, as described by the Minkowski metric. The Minkowski metric can be rewritten as a summation of velocities and as an apportionment of energy equivalence.

The paper then shows the Schwarzschild metric, which adds a spherical non-rotating mass to the spacetime defined by the Minkowski metric, can also be rewritten as a summation of velocities and as an apportionment of energy equivalence. The apportionment of energy equivalence includes a gravitational component. This indicates gravitational energy has a source and a limit to its magnitude.

2 The Minkowski Metric

The Minkowski metric was originally derived based on Hermann Minkowski's fundamental axiom for space-time set out in an address [3] given in September 1908:

> *The substance at any world-point may always, with the appropriate determination of space and time, be looked upon as at rest.*

Minkowski's fundamental axiom for the space-time continuum indicates that for the substance at a world point (e.g., a particle) there exists a local reference frame, with its own local space and time coordinates, in which the substance is at rest with respect to the local space coordinates (but not with respect to the local time coordinate).

For example, assume the local reference frame for a particle has the local space coordinates (ξ, η, ς) and the local time coordinate τ. For the particle, with respect to the local reference frame,

$$\frac{d\xi}{d\tau} = \frac{d\eta}{d\tau} = \frac{d\varsigma}{d\tau} = 0. \tag{1}$$

The Minkowski metric is often expressed using Cartesian reference coordinates (x, y, z, t) and the local time coordinate τ, i.e.,

$$c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2. \tag{2}$$

The Minkowski metric can also be expressed using spherical coordinates, i.e.,

$$c^2 d\tau^2 = c^2 dt^2 - dr^2 - r^2 d\theta^2 - (r sin\theta)^2 d\phi^2. \tag{3}$$

3 Selection of a reference frame from which to measure velocity

In order to measure velocity in the Minkowski metric (and the Schwarzschild metric) it is important to select and consistently use a reference frame. In the Minkowski metric there are two reference frames to choose from. The first is the local reference frame defined by local coordinates $(\xi, \eta, \varsigma, \tau)$. The other is the reference frame (referred to herein as the coordinate reference frame) defined by reference coordinates (x, y, z, t).

There is a distinct disadvantage to use of the local reference frame to make measurements: in its own local reference frame an object is always at rest, that is, as indicated by (1) there is no spatial velocity, i.e., no change in the values of the local space coordinates (ξ, η, ς) with respect to passage of time as measured by the time coordinate τ.

In the coordinate reference frame, however, there can be a detectable motion through the space coordinates. This is referred to herein as spatial velocity $(\vec{v}_s)$, which is a vector sum of the motion in three dimensions of space, i.e.,

$$\vec{v}_s = \vec{v}_x + \vec{v}_y + \vec{v}_z, \tag{4}$$

and which has a magnitude v_s where

$$v_s = \left|\vec{v}_s\right| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}, \tag{5}$$

as measured by the coordinate reference frame.

Because of this distinct advantage of making measurements from the coordinate reference frame, this is the reference frame that will be consistently used herein to make measurements.

4 Expressing the Minkowski Metric as a sum of velocities

The Minkowski metric, shown in (2), can be rearranged into the form of a sum of velocities. Since the observer is making measurements from the coordinate reference frame, momentum and energy will need to be measured with respect to changes in the reference time coordinate t. The Minkowski metric is therefore rearranged to show this. Specifically, (2) can be rearranged as

$$c^2dt^2 = c^2d\tau^2 + dx^2 + dy^2 + dz^2, \tag{6}$$

and therefore,

$$c^2 = c^2\left(\frac{d\tau}{dt}\right)^2 + \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2, \tag{7}$$

which can be reduced to

$$c^2 = c^2\left(\frac{d\tau}{dt}\right)^2 + v_s^2. \tag{8}$$

Let a time velocity v_τ be defined as

$$v_\tau = c\frac{d\tau}{dt}, \tag{9}$$

so that v_τ is a measure of the rate of passage of time as measured by the local time coordinate τ with respect to the rate of the passage of time as measured by the reference time coordinate t. This allows (7) to be rewritten as

$$c^2 = v_\tau^2 + v_s^2. \tag{10}$$

Since the time dimension is regarded as being orthogonal to the space dimensions, (10) can be written in the form of a vector sum, i.e.,

$$c = \left|\vec{v}_\tau + \vec{v}_s\right|. \tag{11}$$

Equation (11) is the Minkowski metric written as a sum of velocities. That is, the vector sum of the velocity in the dimensions of time and space is always equal to the speed of light c.

5 Energy equivalence in the Minkowski metric

The Minkowski metric, like all solutions to Einstein's field equations, describes a matterless field [1, p. 143]. In order to see how the Minkowski metric apportions energy equivalence, it is only necessary to place a particle with mass m anywhere in the field. From (11), a momentum of mass m across four dimensions of time and space can be expressed as

$$mc = \left|m\vec{v}_\tau + m\vec{v}_s\right|. \tag{12}$$

Equation (10) can also be rewritten as

$$mc^2 = mv_\tau^2 + mv_s^2. \tag{13}$$

Equation (13) indicates how the Minkowski metric apportions the energy equivalence [4],

$$E = mc^2, \tag{14}$$

of mass m into an energy component E_τ in the time dimension, where

$$E_\tau = mv_\tau^2, \tag{15}$$

and an energy component in the space dimensions, where

$$E_s = mv_s^2, \tag{16}$$

so that

$$E = mc^2 = E_\tau + E_s. \tag{17}$$

6 The Schwarzschild metric

The full Schwarzschild metric for a spherical non-rotating mass M with a Schwarzschild radius R, is typically expressed with the reference coordinates in the form of spherical coordinates, i.e.,

$$c^2d\tau^2 = c^2\left(1-\frac{R}{r}\right)dt^2 - \frac{dr^2}{(1-R/r)} - r^2d\theta^2 - (r\sin\theta)^2d\phi^2. \tag{18}$$

When $M = 0$ and thus $R = 0$, the Schwarzschild metric reduces to the Minkowski metric.

7 Expressing the Schwarzschild Metric as a sum of velocities

In order to express the Schwarzschild metric as a sum of velocities, a gravitational velocity v_g can be defined using the Newtonian definition of gravitational escape velocity, that is

$$v_g = c\sqrt{\frac{R}{r}}. \tag{19}$$

Likewise because in the Schwarzschild metric space is curved a spatial velocity velocity v_{ss} through curved space can be defined as

$$v_{ss} = \sqrt{\frac{1}{1-R/r}\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2 + (r\sin\theta)^2\left(\frac{d\phi}{dt}\right)^2}. \tag{20}$$

The Schwarzschild metric in (18) can now be expressed as a sum of the velocities v_τ, v_g and v_{ss}. That is, (19) can be rearranged as

$$\begin{aligned} c^2dt^2 = c^2d\tau^2 + c^2\frac{R}{r}dt^2 + \frac{dr^2}{(1-R/r)} + \\ + r^2d\theta^2 + (r\sin\theta)^2d\phi^2, \end{aligned} \tag{21}$$

and thus

$$\begin{aligned} c^2 = c^2\left(\frac{d\tau}{dt}\right)^2 + c^2\frac{R}{r} + \frac{1}{1-R/r}\left(\frac{dr}{dt}\right)^2 + \\ + r^2\left(\frac{d\theta}{dt}\right)^2 + (r\sin\theta)^2\left(\frac{d\phi}{dt}\right)^2. \end{aligned} \tag{22}$$

Using the definition of v_{ss} set out in (20), the definition of v_τ set out in (9) and the definition of v_g set out in (19), allows (22) to be simplified to

$$c^2 = v_\tau^2 + v_g^2 + v_{ss}^2. \tag{23}$$

If a gravitational dimension is regarded as being orthogonal to both the dimensions of curved space and the time dimension, (23) can be written in the form of a vector sum, i.e.,

$$c = \left|\vec{v}_\tau + \vec{v}_g + \vec{v}_{ss}\right|. \tag{24}$$

Equation (24) is the Schwarzschild metric written as a sum of velocities. That is, the vector sum of the velocity in the dimensions of time, space and gravity is always equal to the speed of light c.

8 Using the Schwarzschild metric to apportion energy equivalence

In order to see how the Schwarzschild metric apportions energy equivalence, it is only necessary to place a particle with mass m anywhere in the field. From (24), a momentum of mass m across five dimensions of time, space and gravity can be expressed as

$$mc = \left|m\vec{v}_\tau + m\vec{v}_g + m\vec{v}_{ss}\right|. \tag{25}$$

Equation (10) can also be rewritten as

$$mc^2 = mv_\tau^2 + mv_g^2 + mv_{ss}^2. \tag{26}$$

Equation (26) indicates how the Schwarzschild metric apportions the energy equivalence of mass m into an energy component E_τ, an energy component E_{ss} in the space dimensions, and an energy E_g component where

$$E_g = mv_g^2, \tag{27}$$

so that

$$E = mc^2 = E_\tau + E_g + E_{ss}. \tag{28}$$

9 Reciprocity in the apportionment of energy equivalence

In a system of two particles, one particle having a mass m_1 and a Schwarzschild radius of R_1 and the other particle having a mass m_2 and a Schwarzschild radius of R_2, the Schwarzschild metric allows the energy equivalence of each mass to be apportion into, time, space and gravity components. For example, when spatial coordinates (r_1, θ_1, ϕ_1) are measured with respect to m_1 and local time τ_1 is measured at the location of m_2, the energy equivalence of m_2 can be apportioned using the Schwarzchild metric,

$$c^2 d\tau_1^2 = c^2\left(1 - \frac{R_1}{r_1}\right)dt_1^2 - \frac{dr_1^2}{(1 - R_1/r_1)} - r_1^2 d\theta_1^2 - (r_1 \sin\theta_1)^2 d\phi_1^2, \tag{29}$$

into the following apportionment of energy equivalence:

$$m_2 c^2 = m_2 v_{\tau_1}^2 + m_2 v_{g_1}^2 + m_2 v_{ss_1}^2 = E_{\tau_1} + E_{g_1} + E_{s_1}. \tag{30}$$

Likewise, when spatial coordinates (r_2, θ_2, ϕ_2) are measured with respect to m_2 and local time τ_2 is measured at the location of m_1, the energy equivalence of m_1 can be apportioned using the Schwarzschild metric,

$$c^2 d\tau_2^2 = c^2\left(1 - \frac{R_2}{r_2}\right)dt_2^2 - \frac{dr_2^2}{(1 - R_2/r_2)} - r_2^2 d\theta_2^2 - (r_2 \sin\theta_2)^2 d\phi_2^2, \tag{31}$$

into the following apportionment of energy equivalence:

$$m_1 c^2 = m_1 v_{\tau_2}^2 + m_1 v_{g_2}^2 + m_1 v_{ss_2}^2 = E_{\tau_2} + E_{g_2} + E_{s_2}. \tag{32}$$

10 Implications

The Schwarzschild metric apportions the energy equivalence of a mass into a time component, a spatial component and a gravitational component. This suggests that the source of gravitational energy is the energy equivalence of the mass affected by gravity and therefore that the magnitude of gravitational energy cannot exceed the energy equivalence of that mass. As pointed out by Weller [5, 6], this presents a very significant difficulty for those who view gravity as an unlimited source of energy to perform such tasks as forming black holes and creating universes. This also tends to confirm the assertions of Schwarzschild [7] and Einstein [8] that there is indeed a maximum density of matter.

Submitted on June 13, 2011 / Accepted on June 21, 2011

References

1. Einstein A. The foundation of the general theory of relativity. *The Principle of Relativity*, Dover Publications, New York, 1923, pp. 111–164.
2. Schwarzschild K. On the gravitational field of a mass point according to Einstein's theory, translated by S. Antoci, A. Loinger. 1999, arXiv:Physics/9905030.
3. Minkowski H. A translation of an address delivered at the 80th Assembly of German Natural Scientists and Physicians, at Cologne. The Principle of Relativity, Dover Publications, New York, 1923, pp. 75–91.
4. Einstein A. Does the inertia of a body depend upon its energy-content? The Principle of Relativity, Dover Publications, New York, 1923, pp. 69–71.
5. Weller D. How black holes violate the conservation of energy. *Progress in Physics*, 2011, v. 1, 87–90.
6. Weller D. Five fallacies used to link black holes to Einstein's relativistic space-time. *Progress in Physics*, 2011, v. 1, 91–95.
7. Schwarzschild K. On the gravitational field of a sphere of incompressible fluid according to Einstein's theory. *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Phys.-Math. Klasse*, 1916, 424 (see also in arXiv: physics/9912033).
8. Einstein A. On a stationary system with spherical symmetry consisting of many gravitating masses. *Annals of Mathematics*, 1939, v. 40, no. 4, 922–936.

A Note on the Quantization Mechanism within the Cold Big Bang Cosmology

Armando V. D. B. Assis
Departamento de Física, Universidade Federal de Santa Catarina — UFSC, Trindade 88040-900, Florianópolis, SC, Brazil.
E-mail: armando.assis@pgfsc.ufsc.br

In my paper [3], I obtain a Cold Big Bang Cosmology, fitting the cosmological data, with an absolute zero primordial temperature, a natural cutoff for the cosmological data to a vanishingly small entropy at a singular microstate of a comoving domain of the cosmological fluid. This solution resides on a negative pressure solution from the general relativity field equation and on a postulate regarding a Heisenberg indeterminacy mechanism related to the energy fluctuation obtained from the solution of the field equations under the Robertson-Walker comoving elementar line element context in virtue of the adoption of the Cosmological Principle. In this paper, we see the, positive, differential energy fluctuation, purely obtained from the general relativity cosmological solution in [3], leads to the quantum mechanical argument of the postulate in [3], provided this energy fluctuation is quantized, strongly supporting the postulate in [3]. I discuss the postulate in [3], showing the result for the energy fluctuation follows from a discreteness hypothesis.

1 To the Heisenberg Indeterminacy Relation

Recalling the eqn. (53) in [3], purely derived from the general relativity field equations under the cosmological context:

$$\delta E_\rho = \frac{E_0^+}{\sqrt{1-\dot{R}^2/c^2}}\frac{\dot{R}\,\delta\dot{R}}{c^2}, \tag{1}$$

the δE_ρ, given by the eqn. (1), seems to be exclusively valid when $\delta\dot{R}$ is infinitesimal, since this expression is a first order expansion term, where we do tacitly suppose the vanishing of high order terms. But its form will remain valid in a case of finite variation, as derived is this paper, under the same conditions presented in [3]. The eqn. (1), in terms of indeterminacy, says:

- There is an indeterminacy δE_ρ, at a given t, hence at a given $R(t)$ and $\dot{R}(t)$, related to a *small* inteterminacy $\delta\dot{R}(t)$.

A given spherical shell within a t-sliced hypersurface of simultaneity must enclose the following indeterminacy, if the least possible infinitesimal continuous variation given by the field equations in [3], eqn.(1) here, presents discreteness, viz., if the δE_ρ cannot be an infinitesimal in its entire meaning, albeit mantaining its very small value, as a vanishingly small quantity, but reaching a minimum, reaching a discrete quantum of energy fluctuation,

$$\sum_{l=1}^{k}\left(\delta E_\rho\right)_l = \left.\frac{E_0^+\dot{R}/c^2}{\sqrt{1-\dot{R}^2/c^2}}\right|_t \sum_{l=1}^{k}\left(\delta\dot{R}\right)_l. \tag{2}$$

The eqn. (2) is obtained from eqn. (1) by the summation over the simultaneous fluctuations within the spherical shell (since the quantum minimal energy is a spatially localized object, and the t-sliced spherical shell, a $R(t)$-spherical subset of simultaneous cosmological points pertaining to a t-sliced hypersurface of simultaneity, is full of cosmological substratum), where k denotes a partition, k fundamental fluctuating pieces of the simultaneous spacelike spherical shell within a t-sliced hypersurface. This sum gives the entire fluctuation within the shell. Since these pieces are within a hypersurface of simulteneity, they have got the same cosmological instant t. Hence, they have the same $R(t)$ and the same $\dot{R}(t)$ (points within the t-sliced spherical shell cannot have different $R(t)$, since $R(t)$ is a one-to-one function $R(t) : t \mapsto R(t)$, and does not depend on spacelike variables; the t-sliced spherical shell is a set of instantaneous points pertaining to a t-sliced hypersurface of simultaneity such that these points are spatially distributed over an t-instantaneous volume enclosed by a t-instantaneous spherical surface with radius $R(t)$), the reason why the summation index l does not take into account the common factor at the right-hand side of the eqn. (2). From eqn. (57) in [3], we rewrite the eqn. (2):

$$\sum_{l=1}^{k}\left(\delta E_\rho\right)_l = \left.\frac{E_0^+R_0^2}{R^3\sqrt{1-\dot{R}^2/c^2}}\right|_t \sum_{l=1}^{k}\left(\delta R\right)_l. \tag{3}$$

Now, we reach the total instantaneous fluctuations within the spherical shell at the cosmological instant t, a sum of spacelike localized instantaneous fundamental fluctuations within the spherical shell, giving the total instantaneous fuctuation within this shell. Being the instantaneous spherical shell full of cosmological fluid at t, at each fundamental position within the spherical shell we have got a fundamental energy fluctuation with its intrinsical and fundamental quantum $R_0 = \sqrt{2Gh/c^3}$ of indeterminacy [3], an inherent spherically symmetric indeterminacy at each position within the t-sliced spacelike shell.

Hence, the total fluctuation is now quantized:

$$N_t \delta E_\rho = \left. \frac{E_0^+ R_0^2}{R^3 \sqrt{1 - \dot{R}^2/c^2}} \right|_t N_t R_0, \qquad (4)$$

where N_t is the number of instantaneous fundamental domains, the number of fundamental fluctuations within the instantaneous spherical shell contained within a t-sliced hypersurface of simultaneity. Since R_0 is a fundamental quantum of local indeterminacy, the same R_0 is common to all the instantaneous spacelike points within the shell, the same $(\delta R)_l = R_0$ quantum of fluctuation at its respective point within the t-instantaneous spherical shell contained in a t-sliced surface of simultaneity for all the points in this shell, $\forall\ l$.* But N_t is given by:

$$N_t = \frac{R^3}{R_0^3}. \qquad (5)$$

Using the eqn. (5) in the eqn. (4), we obtain:

$$N_t\, \delta E_\rho = \frac{E_0^+}{\sqrt{1 - \dot{R}^2/c^2}}. \qquad (6)$$

The eqn. (6) gives the total positive fluctuation whitin the t-instantaneous spherical shell, the result used in my postulate in [3]. Furthermore, comparing the eqns. (1) and (6), we see the infinitesimal relation given by the eqn. (1) is valid in the finite fluctuation process given by the eqn. (6), provided $\dot{R}\delta\dot{R} \approx c^2$, a result used in the appendix of [3] to obtain its eqn. (56).

The Heisenberg indeterminacy principle reads, for the entire fluctuation at a given t:

$$\left(N_t\, \delta E_\rho\right) \delta t = \frac{E_0^+ \delta t}{\sqrt{1 - \dot{R}^2/c^2}} \geq \frac{h}{4\pi}. \qquad (7)$$

The increasing smearing out indeterminacy over the cosmological fluid, related to the primordial indeterminacy in virtue of the Universe expansion as postulated in [3]:

- The actual energy content of the universe is a consequence of the increasing indeterminacy of the primordial era. Any origin of a comoving reference frame within the cosmological substratum has an inherent indeterminacy. Hence, the indeterminacy of the energy content of the universe may create the impression that the universe has not enough energy, raising illusions as dark energy and dark matter speculations. In other words, since the original source of energy emerges as an indeterminacy, we postulate this indeterminacy continues being the energy content of the observational universe: $\delta E(t) = E^+(t) = E_0^+ / \sqrt{1 - \dot{R}^2/c^2}$,

follows from the increasing N_t, as one infers from the eqns. (5) and (7).

*See [3]. We are in a context of validity of the Cosmological Principle.

Acknowledgments

A.V. D.B.A is grateful to Y.H.V.H and CNPq for financial support.

Submitted on June 21, 2011 / Accepted on June 21, 2011

References

1. Bondi H. Cosmology. Dover Publications, Inc., New York, 2010.
2. Bondi H. Negative mass in General Relativity. *Review of Modern Physics*, 1957, v. 29 no. 3, 423–428.
3. Assis A. V. D. B. On the Cold Big Bang Cosmology. *Progress in Physics*, 2011, v. 2 58–63.
4. Carrol S. Spacetime and Geometry. An Introduction to General Relativity. Addison Wesley, San Francisco, 2004.

Comments on the Statistical Nature and on the Irreversibility of the Wave Function Collapse

Armando V.D.B. Assis

Departamento de Física, Universidade Federal de Santa Catarina — UFSC, Trindade 88040-900, Florianópolis, SC, Brazil.
E-mail: armando.assis@pgfsc.ufsc.br

In a previous preprint, [2], reproduced here within the appendix in its revised version, we were confronted, to reach the validity of the second law of thermodynamics for an unique collapse of an unique quantum object, to the necessity of an ensemble of measures to be accomplished within copies of identical isolated systems. The validity of the second law of thermodynamics within the context of the wave function collapse was sustained by the large number of microstates related to a given collapsed state. Now, we will consider just one pure initial state containing just one initial state of the quantum subsystem, not an ensemble of identically prepared initial quantum subsystems, e.g., just one photon from a very low intensity beam prepared with an equiprobable eigenset containing two elements, an unique observation raising two likelihood outcomes. Again, we will show the statistical interpretation must prevail, albeit the quantum subsystem being a singular, unique, pure state element within its unitary quantum subsystem ensemble set. This feature leads to an inherent probabilistic character, even for a pure one-element quantum subsystem object.

1 A toy: the fair coin eigenset

Let a two-state coin, fifty-fifty, with eigenset $\{\phi_1, \phi_2\}$, be our quantum subsystem. The initial state of this unique subsystem reads:

$$\Psi = \sum_{k=1}^{2} a_k \, \phi_k = \frac{\sqrt{2}}{2} \, \phi_1 + \frac{\sqrt{2}}{2} \, \phi_2 \, , \tag{1}$$

with:

$$a_k = \int_V \phi_k^* \Psi \, dV = \frac{\sqrt{2}}{2} \;\; \forall \;\; k \in \{1, 2\} \, . \tag{2}$$

The eigenstates ϕ_1 and ϕ_2 are different eigenstates.

The unique element [given by eqn. (1)] subsystem plus an unique ideal apparatus subsystem Φ will define an isolated system. The memory state of the subsystem apparatus is initially empty, and the initial state of the system is:

$$\Psi\Phi|_{t=0} = \left(\frac{\sqrt{2}}{2} \, \phi_1 + \frac{\sqrt{2}}{2} \, \phi_2 \right) \Phi_{[\text{void}]} . \tag{3}$$

After a measure operator U acting on $\Psi \otimes \Phi|_{t=0}$, the system propagates forward in time to the $(t = \tau)$-state, the collapsed state for short:

$$\Psi\Phi|_{t=\tau} = \frac{\sqrt{2}}{2} \, \phi_1 \, \phi_{[\phi_1]} + \frac{\sqrt{2}}{2} \, \phi_2 \, \phi_{[\phi_2]} \, . \tag{4}$$

The observer is represented by the Φ apparatus subsystem, being in its own Hilbert state space H_Φ. Since $\Phi_{[\phi_1]}$ and $\Phi_{[\phi_2]}$ are different states belonging to H_Φ, these apparatus states are mutually exclusive in H_Φ.

- How many final microstates of the isolated system are there?

The answer depends on which space the apparatus Φ resides. For Φ, the collapsed microstate is a member of H_Φ. The state given by the eqn. (4) cannot be observed in H_Φ, hence cannot be counted from H_Φ by the apparatus subsystem. There are two possible final states for the hypothetical one-element measure that are members of H_Φ, $\Phi_{[\phi_1]}$ and $\Phi_{[\phi_2]}$, but both cannot be obtained at the same time. The collapse evolves but just one member of H_Φ subsists as an equilibrium apparatus subsystem state after the collapse. The entropy of a final collapsed state $\Phi_{[\phi_k]}$ in $H_\Phi(\tau)$ is zero, since, under an one-element measure with an unique initial quantum coin state given by eqn. (1), there is just one manner to obtain the $\Phi_{[\phi_k]}$ collapsed state, since the other equally like manner leads to a different collapsed state and should not be considered as being another microstate of the same $\Phi_{[\phi_k]}$. But both the possible collapsed states leads to a same final null entropy. This unique object measure leads to reversible collapse, since the variation of entropy between states is null in any case. We will see this unique object quantum subsystem must be related to a global statistical context.

Choosing an unique coin to accomplish an unique measure, one is establishing there exists just an unique way to obtain the initial coin, to construct the initial coin. But, in fact, there is not. You may make the same coin with another bunch of metallic atoms. We do not take it into account, since a set of identical elements is an unitary set, being irrelevant which element we use to accomplish the measure. But two distinct but identical coins do not necessarily lead to identical outcomes. Hence, if one takes into account the identical man-

ners, including the previous global context within the Universe, from which the system may evolve to the collapse, one does not modify the initial null entropy of the system, since identical coins are identical coins into the input ($t = 0$) but not necessarily identical coins from the output ($t = \tau$). Suppose you may construct the unique coin only from two different ways, $\mathbb{W}_1$ or $\mathbb{W}_2$. Via $\mathbb{W}_1$, there is one possible microstate for each collapsed result as observed by $\Phi_{[\phi_1]}$ *or* $\Phi_{[\phi_2]}$ in the apparatus subsystem reality. In the apparatus reality, the initial number of microstates of the system is also vanished, since the initial number of microstates is 1×1 (in the apparatus world, we do not describe the system via eqn. (3), since this is an object that is not an element of H_Φ. The initial state of the quantum subsystem, our coin, given by the eqn. (1), is unique for $\Phi_{[\text{void}]}$. Initially, there is just one possibility for each subsystem state in the apparatus reality, hence $w_0 = 1 \times 1$ is the initial number of microstates of the [global] system as observed within the apparatus reality. The apparatus dialectics does not handle objects like the ones in the eqns. (3) and (4).). Analogously, via $\mathbb{W}_2$, there is one possible microstate for each collapsed state. But, when $\mathbb{W}_1$ and $\mathbb{W}_2$ are taken into consideration, two possible microstates emerge for each collapsed state, with the same initial null entropy.

When one accomplishes an one-element collapse experiment with various identical initial quantum subsystems (e.g., taking $\mathbb{W}_1$ and $\mathbb{W}_2$ into consideration), the result is one between the possible ones from identical objects (indistinguishable coins). A particular collapse result turns out to be inserted in a global probabilistic context related to the various identical manners by which the Universe may evolve from the past to their states in which there exist identical isolated experiments to be initiated at $t = 0$, in virtue of the entropic evolution of the Universe. The Universe entropically evolves under their various possibilities, and two different manners to construct a same coin are different ways under which the Universe may evolve to a same initial coin state, hence the null entropy, but not necessarily to the same collapsed state. Hence, even an isolated collapse from an unique coin has a global statistical context related to the different manners the Universe might have evolved, and an unique coin exhibits its global statistical bias. Since the Universe is large, a given initial subsystem, our two-state coin initial quantum subsystem, has a miriad of possible histories up to $t = 0$, say N_1, but with none of these manners giving a different object, all giving the same Ψ at $t = 0$. Analogously, one has, as $\Phi_{[\text{void}]}$ possible initial states, a bag with N_2 identical elements, all given by $\Phi_{[\text{void}]}$. When you isolate the system, you obtain an isolated bag with $N_1 \times N_2$ identical elements given by the eqn. (3). The number of microstates related to this bag is $w_0 = 1$. The number of microstates related to $\Phi_{[\phi_k]}$ is not $w_f = 1$ anymore, but [2]:

$$\lim_{N_1 N_2 \to \infty} \sum_{l=1}^{N_1 N_2} \xi_l^p = \lim_{N_1 N_2 \to \infty} \left[\frac{N_1 N_2}{2} + f(N_1 N_2) \right] > 1, \quad (5)$$

being $N_1 \times N_2$ the number of final histories of collapse, where the histories are, now, being instantaneously counted at τ, within the Universe entropic evolution.

Taking into account the the different manners by which an initial subsystem may be obtained does not change the probability of a given collapsed state, conversely, defines it via a natural frequentist interpretation within a global context. The probability associated to a given collapsed state when an unique experiment is accomplished with an unique one-element initial state is the one associated to the frequentist interpretation taking into account the various manners to construct the initial state. Since the Universe may provide infinitely many manners to construct an isolated system, when one takes an exemplar into account, the probabilistical character is inherent to individual processes, since a particular result resides within a global statistical context related to different states of the Universe that leads to the same initial isolated system. Even a single photon within a low intensity beam may be constructed by different manners. A single photon does not know this, obviously, but it behaves under a global statistical context related to the different manners by which the Universe may evolve to that in which a beam of a single photon is within an isolated system with an apparatus.

There are not two final microstates, $\Phi_{[\phi_1]}$ and $\Phi_{[\phi_2]}$, for the collapsed apparatus, and one should not say the entropy variation is $\Delta S = k \ln 2 - k \ln 1$, since different microstates are physically distinguishable a posteriori, carrying different measurable physical properties, encapsulated within the difference between the eigenvectors $\Phi_{[\phi_1]}$ and $\Phi_{[\phi_2]}$. In fact, an unique one-element collapse is a reversible process for quantum initial subsystems with just an unique element. But it is very difficult to observe, since the Universe entropically evolves among a miriad of possibilities leading to identical initial quantum subsystems, inserting an individual measure within the Universe's entropy evolution statistical context, being the number of final collapsed microstates of a given collapsed state greater than 1, leading to an irreversible collapse even with a single photon beam as initial quantum subsystem, e.g., since this single photon within the beam turns out to be in a context of a very large number of available microstates for each possible collapsed state, a context in which the final entropy of a given collapsed state is greater than the initial null entropy.

2 Appendix: comments on the entropy of the wave function collapse

2.1 The Boltzmann formula: a source of misconception for a reckless vision

At a first glance [1], one may think the wave function collapse violates the second law of thermodynamics, since a quantum system prepared as a superposition of eigenstates of a given operator suddenly undergoes to a more restrictive state. But this is not the case, in virtue of the fact that a superposition

and a eigenstate are states on equal footing. The use of the Boltzmann formula:

$$S = k \ln w, \tag{6}$$

for the entropy S of a thermodynamically closed system leads, at a first glance, to the impression that the entropy should have a greater value before the collapse, under an erroneous assumption that the initial number, w_0, of microstates, w, should be greater than the final number of microstates, w_f, in virtue of the needed quantity of eigenstates, $w_0 > 1$, used to construct the wave function before the collapse, in contrast to the apparent $w_f = 1$ after the collapse, where $k = 1.38 \times 10^{-23}\, JK^{-1}$ is the Boltzmann constant. We will see that the converse occurs. Furthermore, one should, firstly, define the thermodynamically closed system as consisting of two subsystems: the quantum object subsystem plus the classical apparatus subsystem.

2.2 A simple solution for this apparent paradox

Consider a quantum subsystem Ψ [4]: prepared as a superposition of the n eigenstates $\{\phi_k\}$, with $1 \leq k \leq n$, of a given operator Φ with finite non-degenerated spectrum:

$$\Psi = a_1 \phi_1 + \cdots + a_n \phi_n = \sum_{k=1}^{n} a_k \phi_k, \tag{7}$$

where:

$$a_k = \int_V \phi_k^* \Psi dV, \tag{8}$$

is the inner product the Hilbert state space is equipped with. The * denotes the complex conjugation and dV the elementar volume of the physical space V of a given representation.

Up to the measure, before the interaction between a classical apparatus subsystem, designed to obtain observable eingenvalues of the operator Φ, and a quantum subsystem Ψ given by eqn. (7), there exists just one microstate of the global system consisted by apparatus subsystem plus quantum subsystem, since these two subsystems are not initially correlated and the initial microstate of the quantum subsystem Ψ is just the unique state Ψ as well the initial microstate of the classical apparatus subsystem is the unique one in which it has no registered eigenvalue.

Hence, in virtue of the initial independence of the subsystems, the initial microstate of the global thermodynamically closed system has multiplicity $w_0 = 1 \times 1 = 1$, being the initial entropy of the global system given by:

$$S_0 = k \ln 1 = 0, \tag{9}$$

in virtue of the eqn. (6).

One may argue the initial state of the classical apparatus subsystem has got a multiplicity greater than 1, since this subsystem seems to have internal modes compatible with an empty memory. We emphasize this is not the case, since the state of the memory defines the apparatus state, being this state an empty one in spite of any apparatus internal modes before an accomplished measure*. The same comment is valid for the quantum subsystem, since the state of this subsystem is Ψ, previously defined by the superposition of a Φ operator eigenstates, $\{\phi_k\}$, being the object Ψ an unique one. These objects, by definition, are initially constrained to these defined states, and one does not need to take into account the different manners by which these subsystems should equally evolve to their respective initial states.

Once a measure is accomplished, there will exist n possible eigenvalues to be registered within the memory of the classical apparatus subsystem, viz., since there are n different final situations for the global system, where n is the number of non-degenerated eigenvectors of the Φ operator. A reckless short-term analysis would lead to the conclusion that the final number of microstates of the global system, w_f, should be $w_f = n$, since it seems to be the number of ways by which a final collapsed state is reached. But such a conclusion is wrong, since the final state is not simply a collapsed one with a label on it. Differently from a case in which a pair of unbiased dice is thrown, where a particular result of a throw of dice is not physically different from any other result, except for the labels on them, a given collapsed state encapsulates physical content. Each collapsed state is a different final state with its characteristic multiplicity, and one should not enroll the possible collapsed states within a same bag with $w_f = n$ possible collapsed elements. Comparing with the throw of dice case, if you erased the dice numbers, their labels, you could not infer the difference between the results, but the physical content within the collapsed wave function result would lead one to infer the difference between different results, between different outcomes of collapse of Ψ.

- Different physical characteristics imply different outcomes for the wave function collapse and define evolutions from the initial global system to new states of the global system, instead of different configurations for a same final state.

In the throw of dice example, the different outcomes are different configurations of a same final state. If the collapsed

*The irrelevance of the apparatus internal modes compatible with a given apparatus memory state asserts the hypothesis of an unbiased apparatus subsystem. Any result to be measured by the apparatus subsystem must have the same number of equally like apparatus microstates. If some result was related to a different number of apparatus compatible microstates, the results with the maximal number of apparatus compatible microstates would be biased. The collapse should not be caused by apparata biases. In virtue of this hypothesis, one may neglect the apparatus internal modes compatible with a particular apparatus memory state, since the same number of internal modes is common to all the memory states, and the variation of entropy cancels out the same common number (say w_a): $\Delta S = S_f - S_0 = k \ln\left(w_f \times w_a\right) - k \ln(1 \times 1 \times w_a) = k \ln w_f - k \ln(1 \times 1)$, where w_f is the number of microstates of a given final state of the global isolated system in which the apparatus has registered the respective collapsed state, considering the apparatus memory state as its unique degree of freedom.

wave function was a state with n different possible configurations for this same collapsed state, the final number of microstates would be $w_f = n$, but this is not the case.

For the collapsed states, the multiplicities of the possible final results are not necessarily the same, since they depend on the outcome probabilities of their respective eigenvalues. Let p be the label of the eigenvalue with the least reliable ($\neq 0$)* outcome probability. The outcome probability of a given eigenvalue is given by Max Born's rule, from which the least probability, of the p-labeled eigenvalue, is simply given by $a_p^* a_p$, where [see eqn. (8)]:

$$a_p^* a_p = \left| \int_V \phi_p^* \Psi dV \right|^2 \neq 0. \tag{10}$$

Applying a frequentist interpretation for the probability, the least multiplicity of microstates is† $N a_p^* a_p$, where N is the quantity of *state-balls* within an *a posteriori interpreted quantum-subsystem-urn* (we are emphasizing that the interaction with the classical apparatus subsystem permits a classical‡, under the frequentist sense, a posteriori, interpretation of probabilities, since any quantum effects of probabilistic superposition of amplitudes cease after the collapse, permitting a frequentist interpretation via Born's rule). Such a frequentist interpretation requires $N \to \infty$, i.e., infinitely many measures to be accomplished on identical quantum subsystems by the classical apparatus subsystem, but we will back to this point later.

The least final entropy of the global system, related to the outcome probability of the p-labeled eigenvalue, reads:

$$S_f = k \ln \left(N a_p^* a_p \right). \tag{11}$$

From the eqns. (9) and (11), the least possible entropy variation turns out to be:

$$\Delta S = S_f - S_0 = k \ln \left(N a_p^* a_p \right). \tag{12}$$

From the eqn. (12), we infer that the second law of thermodynamics holds *iff*:

$$N a_p^* a_p \geq 1 \Rightarrow a_p^* a_p \geq \frac{1}{N}, \tag{13}$$

since $N > 0$. Now, we will prove the following theorem:

Theorem: *The second law of thermodynamics holds for the wave function collapse under a frequentist interpretation via Max Born's rule and, once accomplished the collapse, the collapse is an irreversible phenomenon.*

Proof: Suppose the converse, i.e., that the second law of thermodynamics does not hold for the wave function collapse under a frequentist interpretation via Max Born's rule. In virtue of eqn. (12), one has:

$$\Delta S = S_f - S_0 = k \ln \left(N a_p^* a_p \right) < 0 \Rightarrow N a_p^* a_p < 1. \tag{14}$$

Since§ $a_p \neq 0$, $N \geq 1/(a_p^* a_p)$ violates the condition stated by the eqn. (14). But $N \to \infty$, in virtue of the frequentist interpretation, hence $N > 1/(a_p^* a_p)$, and the eqn. (14) is an absurd. We conclude the second law of thermodynamics holds within the terms of this theorem. The proof the collapse is an irreversible phenomenon follows as a corollary of this theorem. In fact:

$$N > 1/(a_p^* a_p) \Rightarrow N a_p^* a_p > 1 \therefore$$

$$\Delta S = k \ln \left(N a_p^* a_p \right) > 0, \tag{15}$$

and the collapse of the wave function is an irreversible phenomenon, being $\Delta S > 0$ the entropy variation of the thermodynamically closed system: quantum subsystem plus classical apparatus subsystem. ■

The law of large numbers states the probability of an event p, P_p, is given by the limit:

$$\lim_{N \to \infty} \frac{\sum_{l=1}^N \xi_l^p}{N} = P_p \,, \tag{16}$$

where ξ_l^p assumes the value 1 when the event p occurs, or zero otherwise. If $a_p^* a_p \equiv P_p \neq 0$, the limit must obey:

$$\lim_{N \to \infty} \frac{\sum_{l=1}^N \xi_l^p}{N} = \frac{\lim_{N \to \infty} \sum_{l=1}^N \xi_l^p}{\lim_{N \to \infty} N} \neq 0. \tag{17}$$

From eqn. (17), we conclude $\lim_{N \to \infty} \sum_{l=1}^N \xi_l^p$ cannot be finite, since N grows without limit. Hence:

$$\lim_{N \to \infty} \sum_{l=1}^N \xi_l^p > 1. \tag{18}$$

Particularly, the eqn. (18) gives the number of microstates of the p-labelled eigenstate, proving the above theorem. Rigorously, one should substitute:

$$N \to N + \frac{f(N)}{a_p^* a_p}, \tag{19}$$

*If $a_p = 0$, the respective eigenstate ϕ_p, within the superposition representing Ψ [see eqn. (7)], turns out to be an impossible collapsed state. Such consideration would be totally void, since the final microstate associated to it would never occur, being $\Delta S = k \ln 0 - k \ln 1 = -\infty$ [see eqns. (6) and (9)] a violation of the second law of thermodynamics, in accordance with the impossibility of a final microstate with $a_p = 0$.

†See the discussion leading to the eqns. (19) and (20), regarding the meaning of N.

‡Here, the classical designation resides within the counting process after the collapse. We are not saying the final collapsed state leads to a classical interpretation of the quantum object, we are emphasizing that the dialectics after the collapse to interpret frequency of a given collapsed state is the classical one via Born's rule. One does not count quantum waves, but the discrete signals of a collapsed object. Surely, alluding, e.g., to the double-slit canonical example, the diffraction pattern on the screen has not a discrete counterpart, but the points on the screen, when the intensity of the source is reduced, have and may be counted.

§Remember the reliability defining the p-labeled eigenstate, see eqn. (10) again and its inherent paragraph.

within the above theorem proof, with:

$$\lim_{N\to\infty} \frac{f(N)}{N} = 0. \tag{20}$$

Such choice leads to:

$$\sum_{l=1}^{N} \xi_l^p = N a_p^* a_p = \left(N + \frac{f(N)}{a_p^* a_p}\right) a_p^* a_p = \\ = N\left(a_p^* a_p + \frac{f(N)}{N}\right) \quad \therefore \tag{21}$$

$$\frac{\sum_{l=1}^{N} \xi_l^p}{N} = a_p^* a_p + \frac{f(N)}{N}. \tag{22}$$

Taking the limit $N \to \infty$ in eqn. (22), we recover the law of large numbers. Taking the limit $N \to \infty$ in eqn. (21), one obtains in virtue of the eqn. (18):

$$\lim_{N\to\infty} \sum_{l=1}^{N} \xi_l^p = \lim_{N\to\infty} \left(N + \frac{f(N)}{a_p^* a_p}\right) a_p^* a_p > 1\,, \tag{23}$$

therefore

$$\lim_{N\to\infty} \left(N + \frac{f(N)}{a_p^* a_p}\right) > \frac{1}{a_p^* a_p}\,. \tag{24}$$

Eqn. (24) is the argument used to prove the theorem, as one infers from the eqn. (19).

3 Conclusion

Finally, we conclude the reversible collapse of the wave function is an extremely rare statistical phenomenon. Once a collapse is reached, it is irreversible since there are a miriad of indistinguishable but distinct outcomes that may be equally reached, leading to a large number of Universe microstates with this same collapsed result. Hence, if one seeks to overcome the collapse: there exist fundamental issues to bypass.

Acknowledgments

A.V. D. B. A is grateful to Y. H.V. H and CNPq for financial support.

Submitted on June 23, 2011 / Accepted on June 29, 2011

References

1. Callen H. B. Thermodynamics and an Introduction to Thermostatistics. John Wiley and Sons, Second Edition, 1985.
2. Assis A.V.D.B. Comments on the entropy of the wave function collapse. viXra.org: 1106.0004.
3. Assis A.V.D.B. On the nature of $a_k^* a_k$ and the emergence of the Born's rule. arXiv.org: 1009.1532.
4. Gasiorowicz S. Quantum Physics. John Wiley and Sons, Third Edition, 2003.

Our Mathematical Universe: I. How the Monster Group Dictates All of Physics

Franklin Potter

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

A 4th family b' quark would confirm that our physical Universe *is* mathematical and is discrete at the Planck scale. I explain how the Fischer-Greiss Monster Group dictates the Standard Model of leptons and quarks in discrete 4-D internal symmetry space and, combined with discrete 4-D spacetime, *uniquely* produces the finite group Weyl E_8 x Weyl E_8 = "Weyl" SO(9,1). The Monster's j-invariant function determines mass ratios of the particles in finite binary rotational subgroups of the Standard Model gauge group, dictates Möbius transformations that lead to the conservation laws, and connects interactions to triality, the Leech lattice, and Golay-24 information coding.

1 Introduction

The ultimate idea that our physical Universe *is* mathematical at the fundamental scale has been conjectured for many centuries. In the past, our marginal understanding of the origin of the physical rules of the Universe has been peppered with huge gaps, but today our increased understanding of fundamental particles promises to eliminate most of those gaps to enable us to determine with reasonable certainty whether this conjecture is true or false.

My principal goal is to show that if a 4th quark family exists, the physical rules of the Universe follow directly from mathematical properties dictated by the Fischer-Greiss Monster Group via the Monster's j-invariant function and the Möbius transformation in discrete spacetime, with everything related to the Golay-24 information code for the Leech lattice.

In a series of articles and conference talks beginning in 1992 [1–3] I have been predicting that a 4th quark family with a b' quark at about 80 GeV and a t' quark at about 2600 GeV will be produced at the colliders. Its detection will support these proposals:

1. The Standard Model (SM) of leptons and quarks provides an excellent approximation to the actual *discrete* symmetry groups of these fundamental particles and requires little modification for extension to the Planck scale.
2. There are 3 lepton families and 4 quark families, each family of two states defined by a different finite binary rotational subgroup of the $SU(2)_L$ x $U(1)_Y$ part of the SM gauge group.
3. The leptons are 3-D polyhedral entities, and the quarks are 4-D polytope entities which combine into 3-D colorless hadrons, color being a 4-D property with exact symmetry derived from 4-D rotations.
4. Lepton and quark approximate mass values are determined by the j-invariant function of elliptic modular functions, being related to the above subgroups and Möbius transformations in both discrete lattice spaces and continuous spaces.
5. Both 4-D spacetime and 4-D internal symmetry space are discrete at the Planck scale, and both spaces can be telescoped upwards mathematically by icosians to 8-D spaces that *uniquely* combine into 10-D discrete spacetime with discrete Weyl E_8 x Weyl E_8 symmetry (not the E_8 x E_8 Lie group of superstrings/M-theory).
6. All the above is related to the Fischer-Greiss Monster Group which herein I argue actually dictates all the rules of physics, except perhaps the entropy law.
7. Consequently, our physical Universe *is* mathematical with only one set of rules and physical constants, which eliminates any multiverse with different values.
8. We live in the only possible Universe, the one with 4-D discrete spacetime dictated by the Monster Group and its relation to information coding and the Leech lattice.

My discrete geometrical approach briefly outlined above fits within the realm of the SM, so its past successes should still apply. One simply must "discretize" the SM lagrangian. Even Noether's theorem works in discrete spaces [4] to connect conservation laws to symmetries, the conserved quantity being continuous but periodic.

2 Brief orientation for discreteness

A few years ago a comprehensive review [5] summarized many of the historical mathematical and physical arguments for considering the Universe to be mathematical. Included were the three hypotheses: (1) the External Reality Hypothesis (ERH) — there exists an external physical reality completely independent of us humans; (2) the Mathematical Universe Hypothesis (MUH) — our external physical reality is a mathematical structure; and (3) the Computable Universe Hypothesis (CUH) — the mathematical structure that is our external physical reality is defined by computable functions. Recall that a computable function must be specifiable by a finite number of bits. The mathematical details are in that article.

The ERH is relatively easy to accept, for the universe certainly existed long before we humans came on the scene. The MUH is the conjecture for which I hope the data from the

colliders will help us decide. One assumption here is that Gödel's Incompleteness Theorem is not an impediment, i.e., there is no limit to being able to determine the ultimate source of all the rules of Nature and what these rules actually are.

The most interesting statements [5] regarding challenges to the CUH are "... virtually all historically successful theories of physics violate the CUH ..." and "The main source of CUH violation comes from incorporating the continuum, usually in the form of real or complex numbers, which cannot even comprise the input to a finite computation since they generically require infinitely many bits to specify." To me, therein lies the problem: *continuous spaces*.

In particle physics, we consider two spaces: (1) a continuous spacetime for particle movement such as translations, rotations and Lorentz transformations, and (2) a continuous internal symmetry space at each spacetime point for the local gauge interactions of the Standard Model. In both spaces we have successfully used continuous functions for our descriptions of the behavior of Nature.

My proposed solution to this problem is to consider both spaces to be discrete spaces "hidden" underneath the continuous approximation, as if we do not yet have enough resolution to detect this discreteness. All our successful physics theories are then excellent *effective* theories containing continuous fields and continuous wave function amplitudes in this approximate world.

We will not be entering a strange new world by considering a discrete approach, for we use difference equations, lattice models, and discrete computations to approximate continuum physics all the time in numerical calculations, and the results are quite reliable and amazingly accurate. Therefore, I suggest that a fundamental discreteness at the Plank scale of about 10^{-35} meters is not unreasonable [3].

The possibility that the Monster Group, whose influence looms over all of mathematics, could dictate all of physics was put forth in several of my previous papers and conference talks over the last two decades, but other physicists have conjectured a similar proposal. What the others have not realized is the direct connection in a *discrete* internal symmetry space from the Monster to the lepton and quark states via the j-invariant of elliptic modular functions. In this article, I provide additional essential arguments to establish the hegemony of the Monster Group and I arrive at the conclusions that spacetime is discrete and our Universe *is* mathematical.

3 The Monster and the j-invariant

The very large discrete symmetry group called the Monster group M is a finite simple group because it has only two normal subgroups, the trivial one-dimensional group and the whole group itself. Finite simple groups can be used as building blocks in that any other type of finite group can be constructed from them. The list of all finite simple groups is: (i) the cyclic groups C_p, with p prime, (ii) the alternating groups A_n, n > 4, (iii) 16 infinite families of Lie groups, and (iv) 26 sporadic groups. The smallest sporadic is the Mathieu Group M_{11} of order 7920 discovered in 1861, while the largest sporadic is the Monster M constructed in 1980 with order of about 8 x 10^{53}. The Monster has 194 different irreducible representations, with the smallest irreducible matrix representations of M being in space dimensions 1, 196883, 21296876, and 842609326.

As I explain in the next section, the most direct connection of M with the SM of leptons and quarks is via the j-invariant of elliptic modular functions

$$j(\tau) = q^{-1} + 744 + 196884q + 21493760q^2 + \dots \quad (1)$$

where $q = e^{2i\pi\tau}$ and τ is a ratio for a 2-D lattice that we will define in a later section. I.e., this 2-D lattice approach in our discrete spaces leads directly to the symmetry groups for the lepton and quark families and for the Lorentz transformations in spacetime.

As has been determined by mathematicians, the coefficients of the powers of q are simple linear combinations of dimensions of irreducible representations of the identity operation of M, a correlation known as "Monstrous Moonshine". E.g., 196884 = 1 + 196883, and 21493760 = 21296876 + 196883 + 1, etc. More mathematical and historical information about the Monster can be learned from the online papers and books by T. Gannon [6].

4 Binary rotation groups and the j-invariant

Here I review the connection between the j-invariant and the discrete symmetry groups for the leptons and quarks. I have proposed [1–3] that the lepton and quark flavors, being electroweak eigenstates, correspond to orthogonal states in specific discrete symmetry groups called finite binary rotational groups. These seven subgroups of the SM local gauge group act in the $\mathbb{R}^3$ and $\mathbb{R}^4$ real subspaces of the 2-D unitary space $\mathbb{C}^2$ for $SU(2)_L$ x $U(1)_Y$. In fact, I am using discrete $\mathbb{R}^3$ and $\mathbb{R}^4$.

The lepton families correspond to the 3-D finite binary rotational groups called the binary tetrahedral group 2T, the binary octahedral group 2O, and the binary icosahedral group 2I, also labelled as [3, 3, 2], [4, 3, 2], and [5, 3, 2], respectively, in Table 1. These are groups of discrete symmetry rotations and reflections. Binary here refers to the double cover of the SO(3) rotation group by Spin(3) = SU(2), so these groups are finite subgroups of SU(2) and $SU(2)_L$ x $U(1)_Y$.

Having exhausted the group possibilities in $\mathbb{R}^3$, one moves up one real spatial dimension to $\mathbb{R}^4$ in order to define the quark families, which then correspond to the finite binary rotation groups [3, 3, 3], [4, 3, 3], [3, 4, 3], and [5, 3, 3] of the regular 4-D convex polytopes. One may not need the number of quark families to match the number of lepton families for anomaly cancellation because this geometrical approach defines leptons and quarks as 3-D and 4-D entities, respectively. I.e., the interactions are not among point particles.

Leptons						Quarks					
group	order	family	N	Pred. Mass (MeV)	Emp. Mass (MeV)	**group**	order	family	N	Pred. Mass (GeV)	Emp. Mass (GeV)
					.						.
					.	**[3, 3, 3]**	120	$d^{-1/3}$	1/4	0.011	0.007
					.			$u^{+2/3}$		0.38	0.004
[3, 3, 2]	24	e^-	1	[1]	0.511	**[4, 3, 3]**	384	$s^{-1/3}$	1	0.046	0.2
		ν_e		0?	0.0?			$c^{+2/3}$		[1.5]	1.5
[4, 3, 2]	48	μ^-	108	108	103.5	**[3, 4, 3]**	1152	$b^{-1/3}$	108	[5]	5.0
		ν_μ		0?	0.0?			$t^{+2/3}$		160	171.4
[5, 3, 2]	120	τ^-	1728	1728	1771.0	**[5, 3, 3]**	14400	$b'^{-1/3}$	1728	~ 80	?.?
		ν_τ		0?	0.0?			$t'^{+2/3}$		~ 2600	?.?

Table 1: Lepton and quark families for the binary rotational groups [a, b, c], their j-invariant proportionality constant N, and the predicted mass values for the quarks based upon group-to-group N ratios with the charm quark mass [1.5 GeV] and bottom quark mass [5 GeV] as reference masses for ratios of the "up-like" and "down-like" quark states, respectively. These are the "bare" mass predictions. Drawings with these symmetries are online [3].

Each lepton group represents the *binary* rotational symmetries of familiar 3-D regular polyhedrons, the tetrahedron, the octahedron, and the icosahedron. In terms of two complex variables z_1 and z_2, there are three algebraic equations for each regular polyhedron that remain invariant under the operations of its binary group, corresponding to the complex equations for the vertices, the face centers, and the edge centers. Call these three equations W_1, W_2 and W_3, respectively. F. Klein, in a famous 1884 book [7], reported that these three equations are not independent because they form a mathematical syzygy. He showed that two independent equations W_1 and W_2, say, have a ratio proportional to the j-invariant

$$j(\tau) = \frac{W_1}{NW_2} \tag{2}$$

where N is a specific integer, being 1, 108, and 1728, for the three groups, 2T, 2O, and 2I, respectively. Certain integrals, including a mass integral, for the particle states would involve these N values as important factors.

The four binary rotational groups for the quarks are handled [8] by projecting their physical 4-D polytopes onto the 2-D unitary plane $\mathbb{C}^2$ and realizing that their symmetries lead to the same invariant algebraic equations as for the leptons, with the addition of one other symmetry group syzygy for [3, 3, 3]. The corresponding N values are thus 1/4, 1, 108, 1728.

These N values suggest the pairings of the lepton families to quark families as shown horizontally in Table 1. Notice that these family pairings are different from the traditional ad hoc pairings that are normally suggested for the SM because here there exist fundamental geometrical connections.

5 Particle mass values

The influence of the j-invariant of the Monster continues. In spaces where the j-invariant applies, all rational functions (ratio of two polynomials) are proportional to the j-invariant and invariant under all fractional linear transformations (also called Möbius transformations). For physics purposes, mass of a fundamental particle is proportional to the j-invariant because mass is an invariant under Möbius transformations. Conservation laws in physics can be related to Möbius transformations in both discrete and continuous spaces.

At this stage there is no absolute mass scale, so I must use mass ratios only, selecting a different reference mass value for the "up" states and for the "down" states. For the lepton mass values, we have the N ratios 1:108:1728. Table 1 shows the predicted and the actual values. The patterns of ratios match roughly and they were the clue to considering these binary rotational groups.

Note that without using the reference empirical masses for the ratios, the two predicted states in each family would be degenerate with the same mass. One should form two new orthogonal linear superposition states from these original degenerate states. These states would have different "bare mass" values and would be sensitive to the "vacuum" environment.

For the electroweak interactions, a zero-order approximation to the quark CKM mixing matrix and the lepton PMNS mixing matrix follows directly from the characteristic equations of the 3-D and 4-D symmetries projected to the unitary plane $\mathbb{C}^2$, producing unitary eigenvectors and eigenvalues λ_j = exp[iϵ_j]. The two angles (ϵ_1, ϵ_2) are (π, π) for [3,3,2], (2π/3,

$4\pi/3$) for [4,3,2], ($2\pi/5$, $8\pi/5$) for [5,3,2], ($2\pi/5$, $8\pi/5$) for [3,3,3], ($\pi/3$, π) for [4,3,3], ($\pi/6$, $7\pi/6$) for [3,4,3], and ($\pi/15$, $19\pi/15$) for [5,3,3].

One can define a 3 x 3 unitary matrix [9] and substitute angle difference values for the lepton mixing matrix PMNS and a 3 x 3 quark mixing matrix CKM3, producing

$$PMNS = \begin{pmatrix} 0.5 & 0.866 & \epsilon \\ -0.579 & 0.335 & 0.743 \\ 0.643 & -0.372 & 0.669 \end{pmatrix} \quad (3)$$

$$CKM3 = \begin{pmatrix} 0.978 & 0.208 & \epsilon \\ -0.180 & 0.847 & 0.5 \\ 0.104 & -0.489 & 0.866 \end{pmatrix} \quad (4)$$

with ϵ small. Several of the off-diagonal values in VCKM3 would require higher order corrections in order to better agree with empirically determined values.

A 4 x 4 unitary mixing matrix for our four quark families that brings in c_{34} and s_{34} in the 3rd and 4th rows leads to

$$VCKM4 = \begin{pmatrix} 0.978 & 0.208 & \epsilon_1 & \epsilon_2 \\ -0.180 & 0.847 & 0.5 & \epsilon_3 \\ 0.099 & -0.465 & 0.842 & 0.309 \\ -0.032 & 0.151 & -0.268 & 0.951 \end{pmatrix} \quad (5)$$

with all the ϵ values small. Adjustments can be made by considering higher order corrections.

One should not ignore the fact that a degrees-of-freedom argument would make neutrinos that are zero mass exactly. My two lepton states in each family have 4 d.o.f. total, which can partition into the massive electron state with 3 d.o.f., leaving just 1 d.o.f. for the neutrino state. Thus, the neutrino is massless and can have one helicity state only. Alternately, if both lepton states per family share the 4 d.o.f. equally with 2 d.o.f. each, then these would be two massless states, i.e., possibly two sterile neutrino states. Nature appears to have chosen the unequal split, but sterile neutrinos are still a possibility. As to the quarks, the two 4-D quark family states have a total of 6 d.o.f. to split 3-3, guaranteeing the existence of the two massive quark states per family we measure.

The discovery of the b' quark, probably by the FCNC decay b' $\rightarrow$ b + γ, is the acid test of this geometrical approach toward understanding the SM. There is already some hint in the Fermilab data for this decay but the signal/noise ratio is not good enough. The 4th quark family has recently been in vogue because the baryonic particle-antiparticle asymmetry in the Universe (BAU) can then be explained by CP violation with a new value for the Jarlskog invariant that is about 10^{13} times larger [10] than for only 3 quark families. As far as I know, the b' quark remains a viable possibility.

6 Discrete internal symmetry space

In this geometrical approach, the internal symmetry space is discrete $\mathbb{C}^2$ at the Planck scale. Therefore we must consider the mathematical properties of a 2-D hexagonal lattice (or of a 2-D rectangular lattice) of mathematical nodes either with two real axes $\mathbb{R}^2$, or two complex axes $\mathbb{C}^2$, or two quaternion axes $\mathbb{H}^2$, etc. All its nodes can be represented by integer linear combinations of two complex numbers that we label ω_1 and ω_2 forming a right-handed basis (ω_1, ω_2). We can change these two numbers without changing the lattice by letting

$$\begin{aligned} \omega_1' &= a\omega_1 + b\omega_2 \\ \omega_2' &= c\omega_1 + d\omega_2 \end{aligned} \quad (6)$$

where a, b, c, and d, are integer elements of a 2 x 2 matrix with determinant 1. Such matrices form a symmetry group called the "modular group" SL(2, Z) which is related to elliptic curves. Actually, all that matters is the ratio $\tau = \omega_1/\omega_2$ which defines the τ for the j-invariant in Eq. 1. Since

$$f(\tau) = f\left(\frac{a\tau + b}{c\tau + d}\right), \quad (7)$$

all modular functions $f(\tau)$ on the lattice depend only upon its shape. The j-invariant is such a function, and all other SL(2, Z)-invariant functions are rational functions of j(τ).

Eq. 7 defines the fractional linear transformations, i.e., the Möbius transformations, which are based upon the transformations $\tau \rightarrow 1 + \tau$ and $\tau \rightarrow -1/\tau$ for translations, rotations, etc. In the limit when the node spacing approaches zero, the continuous approximation appears and the Möbius transformations include the continuous symmetry transformations.

7 Geometry of the boson interactions

The 12 bosons of the SM, 8 gluons and 4 EW bosons, operate on the fermion states in a continuous internal symmetry space. For a continuous space one can map the complex plane $\mathbb{C} = \mathbb{R}^2$ and unitary plane $\mathbb{C}^2 = \mathbb{R}^4$ to the 2-D Riemann sphere. Its 2-D surface has no demarcations, thus allowing any small or large rotation. Consequently, the symmetry group for the SM interactions is the continuous gauge group of operations.

In my geometrical approach this internal symmetry space is discrete, so only specific finite rotation groups can produce these boson operations. However, when the internal symmetry space is discrete and particle symmetries are defined by the specific finite binary rotation groups for leptons and quarks, the Riemann surface is tessellated, i.e., composed of identical equilateral triangles, their number uniquely determined by the binary rotation group. Then the number of rotational operations becomes severely restricted and each boson operator must respect the integrity of the symmetry group for the lepton or quark families participating in the interaction.

Geometry provides the important clue. We desire a small group in our discrete space for defining these interactions (i.e., producing the appropriate rotations by the bosons), and we find the binary icosahedral group 2I or [5, 3, 2]. However, there will be some missed operations on the symmetry for the binary octahedral group 2O. But if we take 2I twice, i.e., including its "reciprocal" [5, 3, 2], then we get it all.

In order to appreciate this geometry, quaternion algebra simplifies the game. Recall that the SU(2) matrix representation and the unit quaternion **q** are related by

$$\mathbf{q} = w\mathbf{1} + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \Longleftrightarrow \begin{pmatrix} w + iz & x + iy \\ -x + iy & w - iz \end{pmatrix} \qquad (8)$$

where the **i**, **j**, and **k** are unit imaginaries, the coefficients w, x, y, and z are real, and $w^2+x^2+y^2+z^2 = 1$. We can represent the two orthogonal lepton or quark states in each family by two orthogonal unit quaternions in $\mathbb{C}^2$.

There is also a conjugate plane $\mathbb{C}'^2$ for the antiparticles and its Riemann sphere. The conjugate quaternion is **q'** =w**1** - x**i** - y**j** - z**k**. What we discuss for the particle states works for the antiparticle states, too. Having a conjugate space is very special. Clifford algebra and Bott periodicity dictate that only $\mathbb{R}^4$, $\mathbb{R}^8$, and other real spaces $\mathbb{R}^n$ with dimensions divisible by four have two equivalent conjugate spaces. This specific mathematical property dictates a world with both particle states and their antiparticle states for these dimensions only.

One more mathematical fact. The group $U(1)_Y$ for weak hypercharge Y in $SU(2)_L$ x $U(1)_Y$ has the important role of reducing the symmetry between the two spaces, normal and conjugate, in $\mathbb{R}^4 = \mathbb{C}^2$ from being simply equivalent to their being *gauge* equivalent. The physics consequence is that particles and antiparticles have the same positive mass but all other properties can be opposite sign. Alternately, we can use the 2-element inversion group C_i to accomplish the same distinction as well as to determine the intrinsic parity of the particle states, odd for particles and even for antiparticles.

Furthermore, the use of quaternions for the electroweak operations tells us that the L in $SU(2)_L$, which means left-handed chirality only for the weak interaction, is really dictated by quaternion properties, so that the left-handed physics restriction for the weak interaction in $\mathbb{C}^2$ follows. That is, in the normal unitary plane all unit quaternions have left-handed screw transformations that mix the two orthogonal states and right-handed screw transformations that do not. Put another way, the quaternions transform the two orthogonal flavor states as left-handed doublets and right-handed singlets. For example, in the first lepton family, they are (ν_{eL}, e_L) and (ν_{eR}) and (e_R). In the conjugate unitary plane for antiparticles, the quaternion transformations have the opposite handedness.

Now back to rotating the Riemann sphere. In the simplest electroweak (EW) interactions of a boson with an incoming fermion, the fermion state either remains the same (via γ or Z^0) or changes from the initial state to an orthogonal state (via $W^{\pm}$). As examples, the γ may be the identity and the Z^0 may produce a 4π rotation, while the $W^{\pm}$ operates between different states. The 120 operations of the binary icosahedral group 2I are represented by 120 unit quaternions, and 2I contains almost all the rotation operations needed for the 7 fermion family groups. However, several symmetry operations of 2O would be absent. One needs to add the "reciprocal" binary icosahedral group to include all the operations of 2O, making a grand total of 240 operations. (n.b. One could also consider just the generators to realize the same result.)

Here comes an interesting and unexpected mathematical consequence. The first set of 120 quaternions can be expressed as 120 special unit quaternions known as *icosians* which telescope 4-D discrete-space quaternions up to being 8-D discrete-space octonions to locate points that form a special lattice in $\mathbb{R}^8$ called D_8. The second set of 120 quaternions does the same, forming another D_8 lattice in $\mathbb{R}^8$ by filling the holes in the first D_8 lattice.

The icosians are special unit quaternions q_i which have the mathematical form

$$q_i = (e_1 + e_2\sqrt{5}) + (e_3 + e_4\sqrt{5})i + \\ +(e_5 + e_6\sqrt{5})j + (e_7 + e_8\sqrt{5})k \qquad (9)$$

where the eight e_j are special rational numbers. The important mathematical fact here is that in each pair, such as (e_3 + $e_4\sqrt{5}$), exactly one of the e_j is nonzero. Therefore, even though the icosians are telescoping us up to an 8-D space, their primary importance is that they represent 4-D operations in $\mathbb{R}^4$ even though we can now define identical quaternion operations via octonions in the much larger $\mathbb{R}^8$ space also.

Together, these two D_8 lattices of 120 icosians each combine to form the 240 octonions that define the famous E_8 lattice in $\mathbb{R}^8$. The symmetry group for this E_8 lattice is not the Lie group E_8 but the discrete group Weyl E_8.

Therefore, the operations of the SM occur in discrete 4-D internal symmetry space, but they operate also in the discrete 8-D space because these icosians span both spaces simultaneously.

8 Quark color, gluons, and hadron states

Now I must back up to show that the gluon interactions can occur in $\mathbb{R}^4$ for $SU(3)_C$ even though one normally expects the larger space $\mathbb{C}^3$. Because 4-D rotations are simultaneous rotations in two orthogonal planes, each of the three quark color charges Red, Green, and Blue, can be assigned to the three possible rotation plane pairs [wx, yz], [xy, zw], and [yw, xz], respectively. Actually, because these three 4-D rotation pairs are equivalent and we could have made the color assignments in any order, we learn the mathematical reason for color being an *exact* physical symmetry.

Contained within the above specific icosians are the gluon operations on the color states, but one can use a specific 4 x 4 rotation block matrix R to define the transition from one color state in the 4-D space to another. There are 8 orthogonal gluon matrices in agreement with the 8 gluons of the $SU(3)_C$ gauge group of the SM.

Hadrons are colorless quark combinations, so they occur when the combined resultant 4 x 4 matrices produce no net 4-D rotation, i.e., are the identity matrix. One can show that this

colorless state exists for three combinations of quark states only: (1) the quark-antiquark pair with color and anticolor, (2) three quarks, or (3) three antiquarks, with the appropriate linear combinations of colors or anticolors.

The mathematics itself distinguishes quarks (and baryon number) from leptons: the quarks are 4-D entities and the leptons are 3-D entities, with only the 4-D entities capable of the color interaction because color is an exact symmetry in $\mathbb{R}^4$. Quark *confinement* results because isolated quarks are 4-D entities which cannot exist in a 3-D space, so one can never have an isolated single quark in our 3-D spatial world.

The colorless hadron states, being those special mathematical combinations of quark 4-D entities, are now actually 3-D entities like the lepton states are. That is, the colorless combinations of quarks are 3-D composite particle states because their geometrical intersections define 3-D geometric entities.

Therefore, in my geometrical version of the SM, we have 3-D lepton states, 3-D hadron states, 3-D electroweak boson states, but 4-D quark states and 4-D gluon states. The 4-D quark and gluon states are confined, i.e., they cannot exist as separate entities in our 3-D spatial world, but the 3-D lepton, 3-D hadron, and 3-D electroweak boson states can move through 4-D discrete spacetime with its 3 spatial dimensions.

9 Geometry of discrete 4-D spacetime

Our 3-D particles move in discrete 4-D spacetime. We know that continuous 4-D spacetime has symmetries related to its continuous Lorentz group SO(3,1). For a discrete 4-D spacetime and its Lorentz transformations we need to determine a finite subgroup of SO(3,1) for its discrete symmetry.

A clever mathematical approach to 4-D spacetime was introduced by R. Penrose [11] long ago, who showed how to utilize his "heavenly sphere" to account for Lorentz transformations, etc. This "heavenly sphere" is actually 4-D spacetime (t, x, y, z) mapped onto the Riemann sphere. Consider being in the center of the "heavenly sphere" so that light rays from stars overhead pass through unique points on the unit celestial sphere surrounding you. A Lorentz boost is a conformal transformation of the star locations: the constellations will look distorted because the apparent lengths of the lines connecting the stars will change but the angles between these connecting lines will remain the same.

In our discrete 4-D spacetime we need to tesselate this Riemann surface into identical equilateral triangles and then perform the symmetry transformations of the sphere. But we have already achieved this tesselation earlier with the binary rotation groups when we considered the discrete internal symmetry space mapped to the Riemann sphere, so we know the result. Using the isomorphism SO(3,1) = PSL(2,C), we see [2] that the group mathematics connects the conformal transformations just described to the Möbius group via

$$\text{SO}(3,1) = \text{Möbius group} = \text{PSL}(2,\mathbb{C}), \qquad (10)$$

with the discrete Lorentz transformations of the tessellated Riemann sphere already contained in SO(3,1). Thus, we have a unit quaternion group PSL(2,C) (equivalently, an SU(2) matrix or spinor) representation of the Lorentz transformation.

Therefore, we are back to our discrete symmetries of the binary polyhedral groups because they are finite modular subgroups of the Möbius group PSL(2,C). Therefore, the 240 special quaternions called icosians are now required for discrete Lorentz boosts and discrete rotations in the discrete 4-D spacetime. We obtain a second E_8 lattice in $\mathbb{R}^8$ with symmetry group Weyl E_8.

10 Unification of spacetime and the Standard Model

We can now unite the discrete internal symmetry space operations with the discrete spacetime operations [2]. The direct product of our two Weyl E_8 groups results in a subgroup of the continuous group PSL(2,$\mathbb{O}$), where $\mathbb{O}$ represents all the unit octonions. For the continuous case, PSL(2,$\mathbb{C}$) has become PSL(2,$\mathbb{O}$) = SO(9,1), the Lorentz group in 10-D spacetime. That is, the final combined spacetime is *bigger* than I expected, being isomorphic to a 10-D spacetime instead of an 8-D spacetime.

Applying this result to our discrete case, the combined finite subgroup

$$\text{finite PSL}(2,\mathbb{O}) = \text{finite SO}(9,1), \qquad (11)$$

the finite Lorentz group in discrete 10-D spacetime. The same result, expressed in terms of the direct product of the Weyl E_8 groups is

$$\text{Weyl } E_8 \text{ x Weyl } E_8 = \text{"Weyl"} \text{ SO}(9,1), \qquad (12)$$

a finite subgroup of SO(9,1).

Therefore, the big surprise is that the combination of a 4-D discrete spacetime with a 4-D discrete internal symmetry space creates a *unique* connection to 10-D discrete spacetime, not to an 8-D discrete spacetime. Unlike the situation with continuous spaces, we do not have a 6-D "curled up" internal symmetry space with about 10^{500} possibilities.

The mathematics has dictated a beautiful result: there seems to be *only one way* for our Universe to exist when spacetime is discrete.

11 A physical particle model

Even though the mathematics telescopes us up from $\mathbb{R}^4$ to $\mathbb{R}^8$, we still need a physical model of particles in the discrete 4-D spacetime defining our Universe. The leptons, hadrons, and the electroweak bosons are non-point-like 3-D entities that appear to be point-like particles at our normal size resolution of about 10^{11} times larger than the Planck scale.

Peering in at the Planck scale, however, I expect the discrete 4-D internal symmetry space at each spacetime point to conjoin into the discrete 4-D spacetime. In order to do

so, each particle must *emerge* by "gathering in" nodes of the lattice to make its 3-D or 4-D entity with its correct symmetry. For example, if the particle is an electron, we expect the symmetry of the node collection will be tetrahedral to agree with its [3, 3, 2] symmetry. If the lattice of nodes was originally uniformly spaced in this region of discrete space, then the existence of the electron has distorted this lattice with a decreasing distortion amount for increasing distance from the electron's center.

Note that this geometrical approach assumes that the lattice nodes themselves do not have any *measurable* physical properties. Consequently, we have arrived finally at the end of the hierarchy of physical particles within particles. At this point in the geometrical approach we simply must accept this gathering-of-nodes process because the mathematics dictates this process via graph theory and Kuratowski's theorem.

Kuratowski's theorem is important here because it states that a graph is planar if and only if it does not contain a Kuratowski subgraph K_5 or $K_{3,3}$. For example, if an n-dimensional graph (a lattice of nodes) in a spatial dimension higher that 2-D does not contain a Kuratowski K_5 subgraph, also known as the complete graph of five vertices, then this n-D graph reduces to 2-D.

But the first quark family's binary rotational group [3, 3, 3] symmetry is the rotationally symmetric version of the Kuratowski subgraph K_5. Therefore, at least one quark state of the first quark family is stable as it moves through the lattice, while all other quark families have states that will decay down to [3, 3, 3] quark states. Indeed, the physics agrees with this mathematical prediction.

At the DISCRETE'08 conference in December, 2008 where I tried to present this geometrical approach in my allotted 20 minutes (!), C. Jarlskog asked me an interesting question: Why can't the universe have only quarks and gluons? I.e., a QCD world seems complete by itself. Why complicate the material world with leptons and the electroweak interaction? To which I immediately answered: Kuratowski's Theorem in mathematics does not allow such a world, but I was not encouraged to elaborate with any of the details.

Here is the rest of my argument. If quarks are 4-D entities, most quark states decay because they do not have the structure of K_5 (or $K_{3,3}$), so the initial structure will re-form into two or more new particles. In a universe with only quarks and gluons, a problem arises because gluons change only the color state for a particular quark but cannot change one quark flavor into another. In order to obey Kuratowski's theorem, Nature had no choice but to bring in more particles, notably the leptons and the electroweak interaction bosons. Voilà!

The immortality of the electron with group [3,3,2] seems to depend upon its close geometrical relation to the regular K_5 symmetry group [3, 3, 3]. Of course, the electron could annihilate with its antiparticle (and so can a quark).

At this point one might be concerned about the emergence of fermion particles from the "vacuum" state. In order to account for all the particles in the known Universe, the equivalent of about one new hydrogen atom per cubic meter per 10 million years is required. This process can occur because fermions are represented by spinors, and spinors originate from zero-length vectors. That is, according to E. Cartan, one zero-length vector splits into a spinor and conjugate spinor mathematically. The spinor is the fermion such as an electron and the conjugate spinor is the anti-fermion positron, for example. If their total energy remains zero by adding up all energy forms, then this creation process is viable.

As the electron or any 3-D particle moves through the lattice, I would expect that the particle's lattice distortion effect moves with it, with its previous distortions relaxing back toward being a regular lattice while the oncoming positions become more distorted. Mathematically, the Möbius transformations guarantee the integrity of this movement. That is, for our lattice, the transformation $\tau \rightarrow 1+\tau$ ensures that the movement process is identical everywhere in the lattice. The second Möbius transformation $\tau \rightarrow -1/\tau$ when combined with the one above allows rotations and other linear transformations to occur in the lattice.

This lattice distortion by a particle in 4-D discrete spacetime is the "warping of spacetime" associated with the gravitational interaction proposed by A. Einstein in the general theory of relativity. In this way, gravitation appears to be different from the other fundamental interactions which appear to be more localized.

More details of this particle model, such as the geometry of the gravitational interaction, the origin of the rules of quantum mechanics, the origin of time, and the information coding of the fundamental particles, will be discussed thoroughly in the second paper of this series.

12 Triality, the Leech lattice, and information coding

We know that particle EW interactions can be described in lowest order by the Feynman diagram (Fig. 1) involving three particles with three lines meeting at a point. There can be two fermions interacting with one of the electroweak ($W^{\pm}$, Z^0, or γ) or color (8 gluons) bosons. There can be three gluons interacting. More complicated diagrams can be drawn but they will all be made from combinations of this generic one.

This lowest order Feynman diagram with two fermions and one boson is a mathematical *triality* diagram with the fermions representing spinors and the boson representing a vector Jordan algebra entity. Triality is a relationship between three vector spaces over a field F that are all isomorphic to each other. Thus, the common vector space is isomorphic to $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, or $\mathbb{O}$, i.e., involving spinors in dimensions 1, 2, 4, and 8, respectively [12].

In our 4-D discrete spacetime the fermion states can be represented by quaternions. In fact, Clifford algebra tells us that there will be two quaternion representations in $\mathbb{R}^4$ called the right-handed spinor representation S_4^+ and the left-handed

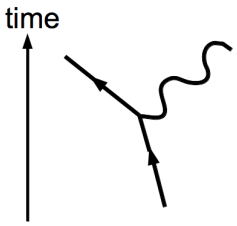


Fig. 1: The incoming fermion emits or absorbs a boson and a fermion exits. E.g., an electron emits a photon and continues in a different direction. This diagram represents triality among two spinors (electron in and electron out) and a vector boson.

spinor representation S_4^-. In general, for even dimensional spacetimes, i.e., even n values, the two spinor representations have dimension $2^{n/2-1}$, but the vector representation has dimension n. For example, in n = 4 space, the boson vector representation is a 4 x 4 real matrix and the fermion spinor representation is a 2 x 2 complex matrix or, equivalently, also a 4 x 4 real matrix. I.e., the fermions and bosons are the same dimension.

We know that the icosians telescope us up to discrete 8-D. With n = 8, the spinor representations are again the same size as the vector representation, both represented by 8 x 8 real matrices. Even so, they are not equivalent representations. However, one can permute the vector, left-handed spinor, and right-handed spinor representations into each other [12]. In 4-D, for example, there is a parity operator that can do this change of a left-handed spinor into a right-handed spinor and vice-versa.

For the generic Feynman diagram, one can think about the two fermions and the one boson as being three E_8 lattices which come together momentarily to form a 24-D lattice called the Leech lattice. The Monster Group again plays its governing role through the j-invariant function. The numerator of $j(\tau)$, being $1 + 720\ q + 146512\ q^2 + \ldots$, is the generating function for the lattice vectors in this product of three copies of the E_8 lattice. And for conformal field theories, the j-invariant is the partition function for the Monster Group [13].

Another very important mathematical connection takes us to information coding theory. One could say that each particle in the triality diagram brings in its 8-bit Hamming code word to temporarily form the 24-bit binary Golay code word or, equivalently, the 12-bit ternary Golay code word, related to the Leech lattice. The 8-bit Hamming code has 72 distinct code words in 9 different but overlapping sets [14], the exact number required for the fundamental particles of the SM: 6 leptons plus 8 x 3 = 24 quarks sums to 30 fermion states; when doubled for anti-particles, makes 60 particle states; then add the 12 bosons to get 72. The 24-bit Golay code word encodes 12 data bits defining up to $2^{12} = 4096$ different items, easily covering the possible interaction triples of the SM.

These code words support the hegemony of the Monster Group because the allowed SM interactions of the leptons and quarks can be related to information theory in 24 dimensions. The second article includes details of the Turyn construction for these Golay-Leech lattice code words and their relationship to quantum information theory and the Monster Group.

13 Conclusion

In this brief article I have outlined specific connections between the mathematics of the Monster Group and fundamental physics particles and rules. These connections support the three hypotheses ERH, MUH, and CUH, so I conclude that the Universe *is* mathematical and that we live in the only possible one. I await the empirical confirmation by the discovery of the 4th quark family, particularly the b' quark at about 80 GeV. Hopefully, the wait will not be long.

Acknowledgements

The author acknowledges Sciencegems.com grant support.

Submitted on June 26, 2011 / Accepted on July 03, 2011

References

1. Potter F. Geometrical Basis for the Standard Model. *International Journal of Theoretical Physics*, 1994, v. 33, 279–305. Online: www.sciencegems.com/gbsm.html
2. Potter F. Unification of Interactions in Discrete Spacetime. *Progress in Physics*, 2006, v. 1, 3–9.
3. Potter F. Discrete Rotational Subgroups of the Standard Model dictate Family Symmetries and Masses. DISCRETE'08 Conference, 2008, Online: www.sciencegems.com/DISCRETE08.PDF
4. Capobianco S. and Toffoli T. Can anything from Noether's theorem be salvaged for discrete dynamical systems? arXiv: 1103.4785v2.
5. Tegmark M. The Mathematical Universe. arXiv: 0704.0646.
6. Gannon T. Postcards from the Edge, or Snapshots of the Theory of Generalised Moonshine. arXiv: math/0109067v1.
7. Klein F. Lectures on the Icosahedron and the Solution (of Equations) of the Fifth Degree. Cosimo Classics, New York, 2007 [originally published 1884].
8. Coxeter H. S. M. Regular Complex Polytopes. Cambridge University Press, Cambridge, 1974.
9. Leontaris G. K. and Vlachos N. D. Knitting neutrino mass textures with or without Tri-Bi maximal mixing. arXiv: 1103.6178v2.
10. Hou W. S. Source of CP Violation for the Baryon Asymmetry of the Universe. arXiv: 1101.2161v1.
11. Penrose R. and Rindler W. Spinors and space-time, Volume 1, Reprint edition. Cambridge University Press, Cambridge, 1987.
12. Baez J. The octonions. arXiv: math.RA/0105155.
13. Witten E. Three-Dimensional Gravity Reconsidered. arXiv: 0706.3359v1.
14. Peng X.-H. and Farrell P. G. On Construction of the (24, 12, 8) Golay Codes. arXiv: cs/0607074v1.

Spin, Isospin and Strong Interaction Dynamics

Eliahu Comay

Charactell Ltd. P.O. Box 39019, Tel Aviv 61390 Israel. E-mail: elicomay@post.tau.ac.il

The structure of spin and isospin is analyzed. Although both spin and isospin are related to the same SU(2) group, they represent different dynamical effects. The Wigner-Racah algebra is used for providing a description of bound states of several Dirac particles in general and of the proton state in particular. Isospin states of the four $\Delta(1232)$ baryons are discussed. The work explains the small contribution of quarks spin to the overall proton spin (the proton spin crisis). It is also proved that the addition of QCD's color is not required for a construction of an antisymmetric state for the $\Delta^{++}(1232)$ baryon.

1 Introduction

The isospin notion has been conceived by W. Heisenberg in 1932 [1, see p. 106]. It aims to construct a mathematical basis that represents the proton-neutron similarity with respect to the strong nuclear force. Both spin and isospin have the same SU(2) group structure. Thus, like spin multiplets of a quantum state, one combines corresponding states of nuclear isobars in an isospin multiplet. For example, the ground state of the ^{14}C, ^{14}O and the $J^{\pi} = 0^{+}$ excited state of ^{14}N are members of an isospin triplet. Obviously, one must remember that isospin is a useful *approximation* that neglects proton-neutron differences that are related to their mass and their electromagnetic interactions.

Later developments have shown that the proton-neutron similarity stems from the similarity between the *u, d* quarks. It follows that the usefulness of isospin symmetry extends to particle physics. For example, the three pions are members of an isospin triplet. Due to historical development, isospin notation takes different form in nuclear and particle physics. Here T and I denote isospin in nuclear and particle physics, respectively. In this work the symbol T is used, mainly because of the following reason. In the case of spin, the symbols J and j denote total and single particle angular momentum operators, respectively. Similarly, the symbols T and t denote the corresponding isospin operators. Thus, due to the same underlying SU(2) group, isospin relations can be readily borrowed from their corresponding spin counterparts. The operators T and t are used in the discussion presented in this work.

This work examines states of electrons and quarks. These particles have spin-1/2 and experimental data are consistent with their elementary pointlike property. Evidently, a theoretical analysis of an elementary pointlike particle is a much simpler task than that of a composite particle. The discussion begins with an examination of relevant properties of electronic states of atoms. The mathematical structure of the SU(2) group is used later for a similar analysis of isospin states.

Two important conclusions are derived from this analysis. First, it is well known that quarks' spin carry only a small fraction of the entire proton's spin [2]. This experimental evidence, which is called the second EMC effect and also the proton spin crisis, is shown here to be an obvious result of the multi-configuration structure of states of more than one Dirac particle. Another result is that the anti-symmetric state of the $\Delta^{++}(1232)$ baryon is well understood and there is no need to introduce a new degree of freedom for its explanation. It means that the historical starting point of the QCD construction has no theoretical basis. (Below, the symbol Δ refers to this isospin quartet of baryons.)

Generally, in order to simplify notation, the specific value of normalization factor is omitted from the expressions. The second and the third sections analyze spin and isospin, respectively. The fourth section provides an explanation for the proton spin crisis. The fifth section explains the antisymmetric structure of the Δ^{++} baryon (without using color). The last section contains concluding remarks.

2 Spin States

A comprehensive discussion of angular momentum can be found in textbooks [3]. In this short work some elements of this theory are mentioned together with a brief explanation. This is done for the purpose of arriving rapidly at the main conclusions. A relativistic notation is used and for this reason the jj coupling [3] takes place.

Let us begin with a discussion of spin and spatial angular momentum. These quantities are dimensionless and this property indicates that they *may* be coupled. Now, the magnetic field depends on space and time. Moreover, the theory must be consistent with the experimental fact where both spatial angular momentum and spin of an electron have the same kind of magnetic field. Thus, it is *required* to construct a relativistically consistent coupling of these quantities. This is the theoretical basis for the well known usage of spin and spatial angular momentum coupling in the analysis of electronic states of atoms.

A motionless free electron is the simplest case and the spin-up electron state is [4, see p. 10]

$$\psi(x^\mu) = Ce^{-imt}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \tag{1}$$

where m denotes the electron's mass.

A second example is the state of an electron bound to a hypothetical pointlike very massive positive charge. Here the electron is bound to a spherically symmetric charge Ze. The general form of a j^π hydrogen atom wave function is [5, see pp. 926–927]

$$\psi(r\theta\phi) = \begin{pmatrix} F\mathcal{Y}_{jlm} \\ G\mathcal{Y}_{jl'm}\end{pmatrix}, \tag{2}$$

where $\mathcal{Y}_{jlm}$ denotes the ordinary Y_{lm} coupled with a spin-1/2 to j, $j = l \pm 1/2$, $l' = l \pm 1$, F, G are radial functions and the parity is $(-1)^l$.

By the general laws of electrodynamics, the state must be an eigenfunction of angular momentum and parity. Furthermore, here we have a problem of *one* electron (the source at the origin is treated as an inert object) and indeed, its wave function (2) is an eigenfunction of both angular momentum and parity [5, see p. 927].

The next problem is a set of n-electrons bound to an attractive positive charge at the origin. (This is a kind of an ideal atom where the source's volume and spin are ignored.) Obviously, the general laws of electrodynamics hold and the system is represented by an eigenfunction of the total angular momentum and parity J^π. Here a single electron is affected by a spherically symmetric attractive field *and* by the repulsive fields of the other electrons. Hence, a single electron does not move in a spherically symmetric field and it *cannot* be represented by a well defined single particle angular momentum and parity.

The general procedure used for solving this problem is to expand the overall state as a sum of configurations. In every configuration, the electrons' single particle angular momentum and parity are well defined. These angular momenta are coupled to the overall angular momentum J and the product of the single particle parity is the parity of the entire system. The role of configurations has already been recognized in the early decades of quantum physics [6]. An application of the first generation of electronic computers has provided a numerical proof of the vital role of finding the correct configuration interaction required for a description of even the simplest case of the ground state of the two electron He atom [7]. The result has proved that several configurations are required for a good description of this state and no configuration dominates the others. This issue plays a very important role in the interpretation of the state of the proton and of the Δ^{++}.

For example, let us write down the 0^+ ground state He_g of the Helium atom as a sum of configurations:

$$\begin{aligned}\psi(\mathrm{He}_g) \;=\; & f_0(r_1)f_0(r_2)_{\frac{1}{2}^+\frac{1}{2}^+} + f_1(r_1)f_1(r_2)_{\frac{1}{2}^-\frac{1}{2}^-} + \\ & f_2(r_1)f_2(r_2)_{\frac{3}{2}^-\frac{3}{2}^-} + f_3(r_1)f_3(r_2)_{\frac{3}{2}^+\frac{3}{2}^+} + \\ & f_4(r_1)f_4(r_2)_{\frac{5}{2}^+\frac{5}{2}^+} + \ldots\end{aligned} \tag{3}$$

Here and below, the radial functions $f_i(r)$, $g_i(r)$ and $h_i(r)$ denote the two-component Dirac radial wave function (multiplied be the corresponding coefficients). In order to couple to $J = 0$ the two single particle j states must be equal and in order to make an even total parity both must have the same parity. These requirements make a severe restriction on acceptable configurations needed for a description of the 0^+ ground state of the He atom.

Higher two-electron total angular momentum allows the usage of a larger number of acceptable configurations. For example, the $J^\pi = 1^-$ state of the He atom can be written as follows:

$$\begin{aligned}\psi(\mathrm{He}_{1^-}) \;=\; & g_0(r_1)h_0(r_2)_{\frac{1}{2}^+\frac{1}{2}^-} + g_1(r_1)h_1(r_2)_{\frac{1}{2}^+\frac{3}{2}^-} + \\ & g_2(r_1)h_2(r_2)_{\frac{1}{2}^-\frac{3}{2}^+} + g_3(r_1)h_3(r_2)_{\frac{3}{2}^-\frac{3}{2}^+} + \\ & g_4(r_1)h_4(r_2)_{\frac{3}{2}^-\frac{5}{2}^+} + g_5(r_1)h_5(r_2)_{\frac{3}{2}^+\frac{5}{2}^-} + \\ & g_6(r_1)h_6(r_2)_{\frac{5}{2}^+\frac{5}{2}^-}\ldots\end{aligned} \tag{4}$$

Using the same rules one can apply simple combinatorial calculations and find a larger number of acceptable configurations for a three or more electron atom. The main conclusion of this section is that, unlike a quite common belief, there are only *three* restrictions on configurations required for a good description of a J^π state of more than one Dirac particles:

1. Each configuration must have the total angular momentum J.
2. Each configuration must have the total parity π.
3. Following the Pauli exclusion principle, each configuration should not contain two or more identical single particle quantum states of the same Dirac particle.

These restrictions indicate that a state can be written as a sum of many configurations, each of which has a well defined single particles angular momentum and parity of its Dirac particles.

The mathematical basis of this procedure is as follows. Take the Hilbert sub-space made of configurations that satisfy the three requirements mentioned above and calculate the Hamiltonian matrix. A diagonalization of this Hamiltonian yields eigenvalues and eigenstates. These eigenvalues and eigenstates are related to a set of physical states that have the given J^π. As pointed out above, calculations show that for a quite good approximation to a quantum state one needs a not very small number of configurations and that no configuration has a dominant weight. These conclusions will be used later in this work.

3 Isopin States

Spin and isospin are based on the same mathematical group called SU(2). Its three generators are denoted j_x, j_y, j_z. An equivalent basis is [1, see pp. 357–363]

$$j_+ = j_x + ij_y, \quad j_- = j_x - ij_y, \quad j_z. \tag{5}$$

All the j operators mentioned above commute with the total j^2 operator. For this reason, if one of them operates on a member of a $(2J + 1)$ multiplet of an SU(2) irreducible representation then the result belongs to this multiplet. The two $j_\pm$ operators are of a particular importance. Thus, let $\psi_{J,M}$ denote a member of such a multiplet and one finds

$$J_z J_- \psi_{J,M} = (M - 1) J_- \psi_{J,M}. \tag{6}$$

This relation means that J_- casts $\psi_{J,M}$ into $\psi_{J,M-1}$

$$J_- \psi_{J,M} = \sqrt{J(J + 1) - M(M - 1)}\, \psi_{J,M-1}, \tag{7}$$

where the appropriate coefficient is written explicitly. Analogous relations hold for the J_+ operator.

Let us turn to isospin. The required operators are simply obtained by taking the mathematical structure of spin and replacing the total spin operator J and the single particle spin operator j by the corresponding isospin operators T, t. (Here, like in the spin case, M, m denote the eigenvalue of T_z, t_z, respectively.) The issue to be examined is the structure of the isospin multiplet of the four baryons:

$$\Delta^-, \Delta^0, \Delta^+, \Delta^{++}. \tag{8}$$

These $\Delta(1232)$ baryons have the lowest energy of the family of the Δ baryons [8]. The Δ^{++} baryon has three u quarks and $\psi_\Delta(uuu)$ denotes its state. Therefore, its isospin state is $T = 3/2$, $M = 3/2$ and the isospin component of the wave function is symmetric with respect to an exchange of any pair of quark.

Let us examine the operation of T_- on Δ^{++}.

$$\begin{aligned} T_- \psi_\Delta(uuu) &= (t_{1-} + t_{2-} + t_{3-}) \psi_\Delta(uuu) \\ &= \psi_\Delta(duu) + \psi_\Delta(udu) + \psi_\Delta(uud), \end{aligned} \tag{9}$$

where t_{i-} operates on the ith quark. This is the way how one obtains a yet unnormalized expression for the Δ^+ baryon from that of Δ^{++}. A successive application of T_- yields expressions for every member of the isospin quartet (8).

Now, the Δ^{++} state is symmetric with respect to its quark constituents and the same symmetry holds for the isospin operator $T_- = t_{1-} + t_{2-} + t_{3-}$. Hence, also the Δ^+ is symmetric with respect to its *uud* quark states. This argument proves that isospin space of *every* member of the baryonic quartet (8) is symmetric. The same result can be obtained from a different argument. The u, d quarks are fermions and their overall state must be antisymmetric with respect to an interchange of any pair of quarks. Now, the isospin operators used above do not affect other coordinates of quarks. It means that for every members of the isospin quartet (8), the entire symmetry of the other coordinates remain antisymmetric and the isospin coordinate is symmetric.

The data confirms the similarity between members of an isospin multiplet. Thus, for example, the mass difference between the Δ^0 and Δ^{++} baryons is less than 3 MeV [8], whereas the mass difference between the Δ multiplet and the nucleons is about 300 MeV. This evidence shows the goodness of the isospin notion, where strong interactions dominate the state of members of an isospin multiplet and the effect of all other interactions can be regarded as a small perturbation.

4 The Proton Spin Crisis

The proton's $J^\pi = 1/2^+$ state is determined by three valence *uud* quarks. The non-negligible probability of the existence of an additional quark-antiquark pair [1, see p. 282] indicates that it is a highly relativistic system. The discussion of section 2 holds for the spin-1/2 point-like quarks and the expansion in configurations is a useful approach. Here the three single particle j^π represent the *uud* quarks, in that order. Evidently, each configuration must satisfy the three requirement written few lines below (4). However, the Pauli exclusion principle of restriction 3 does not hold for the d quark. Thus, in analogy to (3) and (4) one expands the proton's wave function as a sum of terms of specific configurations. A truncated expression for this expansion is shown below:

$$\begin{aligned} \psi(uud) = {} & f_0(r_1) f_0(r_2) h_0(r_3) {}^{\frac{1}{2}^+}_{\frac{1}{2}^+}(0)^{\frac{1}{2}^+} + \\ & f_1(r_1) f_1(r_2) h_1(r_3) {}^{\frac{1}{2}^-}_{\frac{1}{2}^-}(0)^{\frac{1}{2}^+} + \\ & f_2(r_1) g_2(r_2) h_2(r_3) {}^{\frac{1}{2}^+}_{\frac{1}{2}^+}(1)^{\frac{1}{2}^+} + \\ & f_3(r_1) g_3(r_2) h_3(r_3) {}^{\frac{1}{2}^-}_{\frac{1}{2}^-}(1)^{\frac{1}{2}^+} + \\ & f_4(r_1) g_4(r_2) h_4(r_3) {}^{\frac{1}{2}^+}_{\frac{1}{2}^-}(0)^{\frac{1}{2}^-} + \\ & f_5(r_1) g_5(r_2) h_5(r_3) {}^{\frac{1}{2}^+}_{\frac{1}{2}^-}(1)^{\frac{1}{2}^-} + \\ & f_6(r_1) g_6(r_2) h_6(r_3) {}^{\frac{1}{2}^+}_{\frac{3}{2}^+}(1)^{\frac{1}{2}^+} + \\ & f_7(r_1) g_7(r_2) h_7(r_3) {}^{\frac{1}{2}^-}_{\frac{3}{2}^+}(1)^{\frac{1}{2}^-} + \\ & f_8(r_1) g_8(r_2) h_8(r_3) {}^{\frac{1}{2}^+}_{\frac{1}{2}^+}(1)^{\frac{3}{2}^+} + \\ & f_9(r_1) g_9(r_2) h_9(r_3) {}^{\frac{1}{2}^-}_{\frac{1}{2}^-}(1)^{\frac{3}{2}^+} + \\ & f_a(r_1) g_a(r_2) h_a(r_3) {}^{\frac{1}{2}^-}_{\frac{3}{2}^-}(1)^{\frac{1}{2}^+} + \\ & f_b(r_1) g_b(r_2) h_b(r_3) {}^{\frac{1}{2}^+}_{\frac{3}{2}^-}(1)^{\frac{1}{2}^-} + \\ & f_c(r_1) g_c(r_2) h_c(r_3) {}^{\frac{1}{2}^+}_{\frac{1}{2}^-}(1)^{\frac{3}{2}^-} + \ldots \end{aligned} \tag{10}$$

The symbols $_{0\ldots9,a,b,c}$ are used for enumerating the terms of (10). Here, like in (3) and (4), $f_i(r)$, $g_i(r)$ and $h_i(r)$ denote the Dirac two-component radial wave function of the *uud* quarks, respectively (multiplied be the corresponding coefficients). In each term, the number in parentheses indicates how the two angular momenta of the *uu* quarks are coupled. Below, J_{uu} denotes the value of this quantity.

The following remarks explain the form of these terms. An important issue is the coupling of the two *uu* quark that

abide by the Pauli exclusion principle. For this reason, J_{uu} is given explicitly in each term. Another restriction stems from the rule of angular momentum addition. Thus, for every term, the following relation must hold in order to yield a total spin-1/2 for the proton: $J_{uu} = j_d \pm 1/2$. These rules explain the specific structure of each term of (10) which is described below.

In terms 0,1 the two spin-1/2 are coupled antisymmetrically to $J_{uu} = 0$ and the two radial function are the same. In terms 2,3 these spins are coupled symmetrically to $J_{uu} = 1$ and antisymmetry is obtained from the two orthogonal radial functions. In terms 4,5 the different orbitals of the *uu* quarks enable antisymmetrization. Thus, the two spin-1/2 functions are coupled to $J_{uu} = 0$ and $J_{uu} = 1$, respectively. The radial functions are not the same because of the different orbitals. In terms 6,7 the spins are coupled to $J_{uu} = 1$. In terms 8,9 we have a symmetric angular momentum coupling $J_{uu} = 1$ and the antisymmetry is obtained from the orthogonality of the radial function $f_i(r)$, $g_i(r)$. Terms *a,b* are analogous to terms 6,7, respectively. In term *c* the different *uu* orbitals enable antisymmetrization and they are coupled to $J_{uu} = 1$.

A comparison of the expansion of the He atom ground state (3) and that of the proton (10) shows the following points:

1. If the expansion is truncated after the same value of a single particle angular momentum then the number of terms in the proton's expansion is significantly larger.
2. This conclusion is strengthened by the fact that the proton has a non-negligible probability of an additional quark-antiquark pair. Evidently, an inclusion of this pair increases the number of acceptable configurations.
3. Calculations show that the number of configurations required for the ground state spin-0 of the two electron He atom is not very small and that there is no single configuration that dominates the state [7]. Now the proton is a spin-1/2 relativistic particle made of three valence quarks. Therefore, it is very reasonable to assume that its wave function takes a multiconfiguration form.

Using angular momentum algebra, one realizes that in most cases an individual quark does not take the proton's spin direction. This is seen on two levels. First, the upper and the lower parts of the quark single particle function have $l = j \pm 1/2$. Furthermore, the relativistic quark state indicates that the coefficients of the upper and the lower part of the Dirac four component function take a similar size. Hence, for the case where $j = l - 1/2$, the Clebsch-Gordan coefficients [3] used for coupling the spatial angular momentum and the spin indicate that the spin of either the upper or the lower Dirac spinor has no definite direction and that the coefficient of the spin down is not smaller than that of the spin up [3, see p. 519].

Let us turn to the coupling of the quark spins. The 3-quark terms can be divided into two sets having $j_{uu} = 0$ and $j_{uu} > 0$, respectively. For $j_{uu} = 0$ one finds that the single particle $j_d = 1/2$ and this spin is partially parallel to the proton's spin. For cases where $j_{uu} > 0$, the proton's quark spins are coupled in a form where they take both up and down direction so that they practically cancel each other. The additional quark-antiquark pair increases spin direction mixture. It can be concluded that the quark spin contribute a not very large portion of the proton spin and the rest comes from the quark spatial motion. This conclusion is supported by experiment [9].

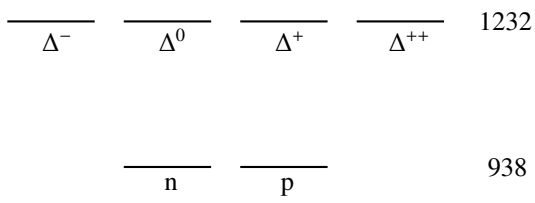


Fig. 1: Energy levels of the nucleon and the Δ isospin multiplets (MeV).

5 The State of the Δ^{++} Baryon

In textbooks it is argued that without QCD, the state of the Δ^{++} baryon demonstrates a fiasco of the Fermi-Dirac statistics [10, see p. 5]. The argument is based on the claim that the Δ^{++} takes the lowest energy state of the Δ baryons [11] and therefore, its spatial wave function consists of three single particle symmetric s-waves of each of its three *uuu* quarks. Now the $J^\pi = 3/2^+$ state of the Δ baryons shows that also their spin is symmetric. It means that the Δ^{++} is regarded to have space, spin and isospin symmetric components of its wave function. As stated above, textbooks claim that this outcome contradicts the Fermi-Dirac statistics. However, using the physical issues discussed in this work and the energy level diagram (see Fig. 1) of the nucleon and the Δ baryons, it is proved that this textbook argument is incorrect.

- As explained in section 3, all members of an isospin multiplet have the same symmetry. Hence, if there is a problem with the Fermi-Dirac statistics of the Δ^{++} then the same problem exists with Δ^+ and Δ^0. It follows that if the above mentioned textbook argument is correct then it is certainly incomplete.
- *The data described in fig. 1 shows that* Δ^+ *is an excited state of the proton.* Hence, its larger mass is completely understood. Thus, there is no problem with the Fermi-Dirac statistics of the Δ^+ baryon. Analogous relations hold for the neutron and the Δ^0 baryons. Using the identical statistical state of the four Δ baryons (8), one realizes that there is no problem with the Fermi-Dirac statistics of the Δ^{++} and the Δ^- baryons.
- The multi-configuration structure of a bound system of

Dirac particles is known for about 50 years [7]. In particular, the multi-configurations structure of all baryons (like in (10)) proves that, contrary to the above mentioned textbook argument [10, see p. 5], the single particle spatial wave functions of the three *u* quarks of the Δ^{++} baryon *are not a pure s-wave*.

6 Conclusions

This work uses the Wigner-Racah mathematical structure and proves two very important points. It explains the small contribution of quark's spin to the overall proton spin. Therefore, it eliminates the basis for the proton spin crisis. It also proves that everything is OK with the Fermi-Dirac statistics of the Δ^{++} baryon. It follows that there is no need to introduce the QCD's color degree of freedom in order to build an antisymmetric wave function for this baryon.

Submitted on July 23, 2011 / Accepted on July 29, 2011

References

1. Perkins D. H. Introductions to High Energy Physics. Addison-Wesley, Menlo Park, 1987.
2. Ashman J. et al. (EMC) A Measurement of the Spin Asymmetry and Determination of the Structure Function g_1 in Deep Inelastic Muon-Proton Scattering. *Physics Letters B*, 1988, v. 206, 364–370.
3. de-Shalit A. and Talmi I. Nuclear Shell Theory. Academic, New York, 1963.
4. Bjorken J. D. and Drell S. D. Relativistic Quantum Mechanics. McGraw, New York, 1964.
5. Messiah A. Quantum Mechanics. Dover, Mineola, 1999.
6. Bethe H. A. Intermediate Quantum Mechanics. Benjamin, New York, 1964.
7. Weiss A. W. Configuration Interaction in Simple Atomic Systems. *Physcal Review*, 1961, v. 122, 1826–1836.
8. Nakamura K. et al. (Particle Data Group), Review of Particle Physics 075021. *Journal of Physics G: Nuclear and Particle Physics*, 2010, v. 37, 1–1422.
9. Kuhn S. E., Chen J. -P., Leader E. Spin structure of the nucleon–status and recent results *Progress in Particle and Nuclear Physics*, 2009, v. 63, 1–50.
10. Halzen F. and Martin A. D. Quarks and Leptons. Wiley, New York, 1984.
11. Today more than 10 different Δ baryonic multiplets are identified [8].

Einstein's Planetary Equation: An Analytical Solution

Chifu E. Ndikilar
Department of Physics, Gombe State University, P.M.B 127, Gombe, Gombe State, Nigeria.
E-mail: ebenechifu@yahoo.com

Einstein's planetary equation can be solved by the method of successive approximations.This yields two linearly independent solutions. An analytical solution is presented for this equation. This solution produces eight linearly independent mathematical solutions, two of which are given approximately by the well-known method of successive approximations.

1 Introduction

Einstein's planetary equation is given [1] by

$$\frac{d^2u}{d\phi^2} + u - \frac{k}{l^2} = \frac{3k}{c^2}u^2 \tag{1}$$

where ϕ and u are the instantaneous angular and reciprocal radial displacements of the planet in the fixed plane of motion, with the Sun as origin, l is the constant angular momentum per unit mass [2] and

$$k = GM \tag{2}$$

where M is the rest mass of the Sun, G is the universal gravitational constant and c is the speed of light in vacuum. The method of successive approximations yields the solution of equation (1) [1] as:

$$r(\phi) = \frac{1}{u(\phi)} = \frac{\left(1-\epsilon_0^2\right)a_0}{1+\epsilon_0\cos\left[\left(1-\frac{3k^2}{c^2l^2}\right)\phi+\alpha\right]} \tag{3}$$

where ϵ_0 is the eccentricity, a_0 the semi-major axis and α is the epoch. The second solution of equation (1) obtained from the method of successive approximations is the solution (3) with sine instead of cosine. The effect revealed by these two approximate solutions is an anomalous precession of the planetary orbit in which the perihelion advances by an angle per revolution Δ given [1] by

$$\Delta = \frac{6\pi k^2}{c^2l^2}. \tag{4}$$

In this article, Einstein's planetary equation (1) is solved analytically.

2 Analytical Solution

Suppose the analytical solution of equation (1) is in the form of a Taylor or Laurent series given as

$$u(\phi) = \sum_{n=0}^{\infty} A_n\, exp\left\{ni\left(\omega\phi + \phi_0\right)\right\} \tag{5}$$

where A_n, ω and ϕ_0 are constants. Then, substituting (5) into (1), applying the linear independence of the exponential functions and equating corresponding coefficients on both sides yields the following system of equations:

$$\frac{3k}{c^2}A_0^2 - A_0 + \frac{k}{l^2} = 0 \tag{6}$$

$$\omega^2 = 1 - \frac{6k}{c^2}A_0 \tag{7}$$

$$A_1 = arbitrary\ constant \tag{8}$$

$$A_2 = \frac{3k}{c^2}\left(1 - 2^2\omega^2 - \frac{6k}{c^2}A_0\right)^{-1}A_1^2 \tag{9}$$

$$\begin{aligned} A_3 \;=\; & \frac{18k^2}{c^4}\left[\left(1 - 2^2\omega^2 - \frac{6k}{c^2}A_0\right)\right. \\ & \left.\left(1 - 3^2\omega^2 - \frac{6k}{c^2}A_0\right)\right]^{-1} A_1^3 \end{aligned} \tag{10}$$

and so on. Equation (6) is a binomial in A_0 and has two possible roots given by

$$A_{0^-} = \frac{c^2}{6k}\left[1 - \left(1 - \frac{12k^2}{c^2l^2}\right)^{1/2}\right] \tag{11}$$

and

$$A_{0^+} = \frac{c^2}{6k}\left[1 + \left(1 - \frac{12k^2}{c^2l^2}\right)^{1/2}\right] \tag{12}$$

It follows from substituting (11) into (7) that they are two possible values of the parameter ω given as:

$$\omega_1 = \left\{1 - \left[1 - \left(1 - \frac{12k^2}{c^2l^2}\right)^{1/2}\right]\right\}^{1/2} \tag{13}$$

and

$$\omega_2 = -\left\{1 - \left[1 - \left(1 - \frac{12k^2}{c^2l^2}\right)^{1/2}\right]\right\}^{1/2} \tag{14}$$

Similarly, by substituting (12) into (7) other two possible values of the parameter are obtained as:

$$\omega_3 = \left\{1 - \left[1 + \left(1 - \frac{12k^2}{c^2l^2}\right)^{1/2}\right]\right\}^{1/2} \tag{15}$$

and

$$\omega_4 = -\left\{1 - \left[1 + \left(1 - \frac{12k^2}{c^2 l^2}\right)^{1/2}\right]\right\}^{1/2}. \quad (16)$$

It follows from equation (9) that A_2 has eight possible values given by

$$A_{2^1} = \frac{3k}{c^2}\left(1 - 2^2\omega_1^2 - \frac{6k}{c^2}A_{0^+}\right)^{-1} A_1^2 \quad (17)$$

$$A_{2^2} = \frac{3k}{c^2}\left(1 - 2^2\omega_1^2 - \frac{6k}{c^2}A_{0^-}\right)^{-1} A_1^2 \quad (18)$$

$$A_{2^3} = \frac{3k}{c^2}\left(1 - 2^2\omega_2^2 - \frac{6k}{c^2}A_{0^+}\right)^{-1} A_1^2 \quad (19)$$

$$A_{2^4} = \frac{3k}{c^2}\left(1 - 2^2\omega_2^2 - \frac{6k}{c^2}A_{0^-}\right)^{-1} A_1^2 \quad (20)$$

$$A_{2^5} = \frac{3k}{c^2}\left(1 - 2^2\omega_3^2 - \frac{6k}{c^2}A_{0^+}\right)^{-1} A_1^2 \quad (21)$$

$$A_{2^6} = \frac{3k}{c^2}\left(1 - 2^2\omega_3^2 - \frac{6k}{c^2}A_{0^-}\right)^{-1} A_1^2 \quad (22)$$

$$A_{2^7} = \frac{3k}{c^2}\left(1 - 2^2\omega_4^2 - \frac{6k}{c^2}A_{0^+}\right)^{-1} A_1^2 \quad (23)$$

$$A_{2^8} = \frac{3k}{c^2}\left(1 - 2^2\omega_4^2 - \frac{6k}{c^2}A_{0^-}\right)^{-1} A_1^2 \quad (24)$$

Similarly, it follows from (10) that A_3 has eight possible values. The above sequence may be continued to derive the eight possible corresponding values for each of the constants A_4, A_5, ... in terms of the arbitrary constant A_1 . This sequence implies eight mathematically possible analytical solutions of Einstein's planetary equation of the form:

$$\begin{aligned} u(\phi) &= A_0 + A_1 exp\left[i\left(\omega\phi + \phi_0\right)\right] + \\ &\quad f_2(A_1) exp\left[2i\left(\omega\phi + \phi_0\right)\right] + ... \\ &\quad f_n exp\left[ni\left(\omega\phi + \phi_0\right)\right] + ... \end{aligned} \quad (25)$$

where ϕ_0 and A_1 are arbitrary.

Now, consider the first exact analytical solution corresponding to equations (12) and (14). In this case, it follows from (9) that

$$A_2 = f_2(A_1) = -\frac{k}{c^2}\left(1 - \frac{6k}{c^2}A_{0^-}\right)^{-1} A_1^2 \quad (26)$$

and

$$A_3 = f_3(A_1) \quad (27)$$

and in general

$$A_n = f_n(A_1),\ n = 4,\ 5,\ ... \quad (28)$$

In this case, the exact analytical solution of Einstein's planetary equation is a complex function of ϕ which may be written in Cartesian form as

$$u(\phi) = x(\phi) + iy(\phi) \quad (29)$$

where

$$\begin{aligned} x(\phi) &= A_{0^-} + A_1\cos\left(\omega_1\phi + \phi_0\right) + \\ &\quad f_2(A_1)\cos 2\left[\left(\omega_1\phi + \phi_0\right)\right] + \ldots \end{aligned} \quad (30)$$

and

$$\begin{aligned} y(\phi) &= A_{0^-} + A_1\sin\left(\omega_1\phi + \phi_0\right) + \\ &\quad f_2(A_1)\sin 2\left[\left(\omega_1\phi + \phi_0\right)\right] + \ldots \end{aligned} \quad (31)$$

Therefore it may be expressed in Euler form as

$$u(\phi) = R(\phi)e^{i\Phi(\phi)} \quad (32)$$

where R is the magnitude given by

$$R(\phi) = \left\{x^2(\phi) + y^2(\phi)\right\}^{\frac{1}{2}} \quad (33)$$

and Φ is the argument given by

$$\Phi(\phi) = \tan^{-1}\left\{\frac{y(\phi)}{x(\phi)}\right\}. \quad (34)$$

Hence by definition the instantaneous radial coordinate of the planet from the Sun, r , is given by

$$r(\phi) = R^{-1}(\phi)\,\ell^{-i\Phi(\phi)}. \quad (35)$$

3 Physical Interpretation of First Analytical Solution

The instantaneous complex radial displacement r of the planet from the Sun is given in terms of the angular displacement Φ as

$$r(\phi) = R^{-1}(\phi)\,\ell^{-i\Phi(\phi)}. \quad (36)$$

Therefore the magnitude of the instantaneous complex radial displacement of the planet from the Sun can be considered to be the real physically measurable instantaneous radial displacement, r_p. Thus,

$$r_p(\phi) = R^{-1}(\phi) = \left\{x^2(\phi) + y^2(\phi)\right\}^{-\frac{1}{2}}. \quad (37)$$

It may be noted from (9) and (10) that for $n > 1$ $f_n(A_1)$ is of order at most c^{-2n}. Therefore as a first approximation let us neglect all terms in $f_n(A_1)$ for $n > 1$. Then it follows from (37) and (31)–(32) that

$$r_p(\phi) = \frac{A}{1 + \varepsilon_1\cos\left(\omega_1\phi + \phi_0\right)} \quad (38)$$

where

$$A = \frac{1}{A_{0^-}}\left(1 + \frac{A_1^2}{A_{0^-}^2}\right)^{-\frac{1}{2}} \quad (39)$$

and

$$\varepsilon_1 = \frac{A_1}{A_{0-}}\left(1 + \frac{A_1^2}{A_{0-}^2}\right)^{-1}. \tag{40}$$

Consequently, the orbit is a precesing conic section with eccentricity and hence semi-major axis given by

$$a = \frac{A}{1 - \varepsilon_1^2} \tag{41}$$

and perihelion displacement angle Δ given by

$$\Delta = 2\pi\left(\omega_1^{-1} - 1\right). \tag{42}$$

It follows from (42) and (14) that the perihelion displacement angle from this analytical method is given explicitly as

$$\Delta = \frac{6\pi k^2}{c^2 l^2} + \frac{54\pi k^4}{c^4 l^4}. \tag{43}$$

This is an advance precisely as obtained from the method of successive approximations. The leading term in (43) is identically the same as the leading term of the corresponding advance from the method of successive approximations [1]. Moreso, this analytical method reveals the exact corrections of all orders of c^{-2} to the leading term in (44).

It also follows from (40) and (12) that the orbital eccentricity ε_1 from this analytical method is given explicitly as

$$\begin{aligned} \varepsilon_1 \;=\; & \frac{l^2 A_1}{k}\left(1 + \frac{3k^2}{c^2 l^2} + \ldots\right)^{-1} \\ & \left[1 + \frac{l^4 A_1^2}{k^2}\left(1 + \frac{3k^2}{c^2 l^2} + \ldots\right)^{-2}\right]^{-1}. \end{aligned} \tag{44}$$

Thus, an experimental measurement of the orbital eccentricity ε_1 in equation (45) is sufficient to determine the parameter A_1 that occurs in the exact analytical solution. It also follows from this result that the analytical method in this article reveals post-Newtonian corrections of all order of c^{-2} to the planetary orbital eccentricity which have not been derived from the method of successive approximations.

It also follows from equations (41) and (14) that the orbital semi-major axis from this analytical method is given explicitly as

$$\begin{aligned} a \;=\; & \frac{l^2}{\left(1 - \varepsilon_1^2\right)k}\left(1 + \frac{3k^2}{c^2 l^2} + \ldots\right)^{-1} \\ & \left[1 + \frac{l^4 A_1^2}{k^2}\left(1 + \frac{3k^2}{c^2 l^2} + \ldots\right)^{-2}\right]^{-1}. \end{aligned} \tag{45}$$

Thus, this analytical method reveals post-Newtonian corrections of all orders of c^{-2} to planetary semi-major axis, which have not been derived from the method of successive approximations.

4 Conclusion

This article uncovers an analytical solution to Einstein's planetary equation. The first analytical solution to the order of c^{-2}, reveals post-Newtonian corrections to the orbital eccentricity and semi-major axis of a planet. Moreover, up to the second iterate there is no such correction from the method of successive approximations. Consequently, these unknown corrections to orbital eccentricity revealed by the analytical approach in this article are opened up for experimental investigation.

Submitted on August 11, 2011 / Accepted on August 18, 2011

References

1. Weinberg S. Gravitation and Applications of the General Theory of Relativity. John Wiley, New York, 1972, pp. 175–210.
2. Ndikilar C. E., Usman A., Meludu O. C. Orbits in Homogeneous Oblate Spheroidal Gravitational Space Time. *Progress in Physics*, 2009, v. 3, 49–53.

Building Galactic Density Profiles

Yuri Heymann

3 rue Chandieu, 1202 Geneva, Switzerland.

E-mail: y.heymann@yahoo.com

The principal objective of this study is to provide a method to build galactic density profiles. The models developed in this study were tested against the zCosmos deep field galactic survey. The herein study suggests that light travel distances need to be converted into Euclidean distances in order to derive the galactic density profile of the survey which is the evolution of galactic density over time. In addition, the present study indicates an Ω_m of 0.19.

1 Introduction

The main purpose of the herein study is to provide a method to build galactic density profiles which requires the conversion of light travel distances (LTD) to Euclidean distances. The LTD is the distance traversed by a photon between the time it is emitted and the time it reaches the observer. In astronomical units, the Euclidean distance is defined as the equivalent distance that would be traversed by a photon between the time it is emitted and the time it reaches the observer if there were no expansion of the Universe.

The zCosmos deep field was used to derive the galatic density profile based on a sampling method, and to compute an estimate of the mean mass density of the Universe.

2 Mathematical development and methods

Galactic density profiles have been derived from the normalization of the galactic counts between redshift buckets by dividing by the corresponding sample volume. For the scenario with additive LTD, the LTDs were directly fed into the sampling volume formula eq. (2). For the scenario with a model of the motion of the photon in an expanding space, the Euclidean distances were fed into the sampling volume formula.

2.1 Method to build galactic density profiles

2.1.1 Normalisation of galactic counts

Let us consider an observer positioned at the center of a sphere of radius r and looking at a cone of sky in the z direction. The observer is counting galaxies within this cone, and measures the redshift for each object. A histogram of the galactic counts versus redshifts is obtained by counting the set of objects contained within each redshift bucket. This histogram is required to be normalised in order to obtain the density profile. Below is derived the expression of the sampling volume of the buckets, function of r_0 the lower radius of the sampling bucket, and Δr the radius width of the bucket. The sampling volume in spherical coordinates is described by the following integral:

$$V_{r_0 o,\Delta r} = \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\theta_0} \sin\theta \, d\theta \, d\varphi \int_{r_0}^{r_o+\Delta r} r^2 dr. \quad (1)$$

By solving integral (1), the sampling volume for a spherical sampling ($\theta_0 = \pi$) is expressed as following:

$$V_{r_0,\Delta r} = \frac{4\pi}{3}\left((r_0+\Delta r)^3 - r_0^3\right), \quad (2)$$

where $V_{r_0,\Delta r}$ is the sampling volume for a given bucket, r_0 the lower radius of the bucket, and Δr the radius width of the bucket.

In order to use eq. (2), the galactic counts need to be converted into spherical values, by multiplying the counts by the sphere to survey solid angle ratio (η). Given the zCosmos survey spectroscopic area of 0.075 square degrees which is the solid angle, this ratio is the following:

$$\eta = \frac{4\pi\,(180/\pi)^2}{0.075} = 550'038. \quad (3)$$

The reported survey coverage area of the zCosmos-deep field is 1 deg2, [8]. However, what is required is the solid angle which is measured by the area of the survey projected in the plan described by the right ascension in degrees and $180/\pi * \sin(declination)$. Note that the sine of declination term is due to the Jacobian for spherical coordinates. The spectroscopic area obtained with this procedure is 0.075 deg2 (surface coverage in figure 1).

2.1.2 Conversion of redshifts to LTDs

Two approaches are available for converting the redshifts from observed galaxies into LTDs, one based on cosmological redshifts and the other one on dopplerian redshifts. First, let us introduce the method based on cosmological redshifts from the calculator of Wright [16] which uses a Lambda-CDM cosmology. The followings are generally assumed for this model: a flat Universe, with parameters: Ω_M = 0.27, Ω_{vac} = 0.73 and H_o = 71 [km s^{-1} Mpc^{-1}].

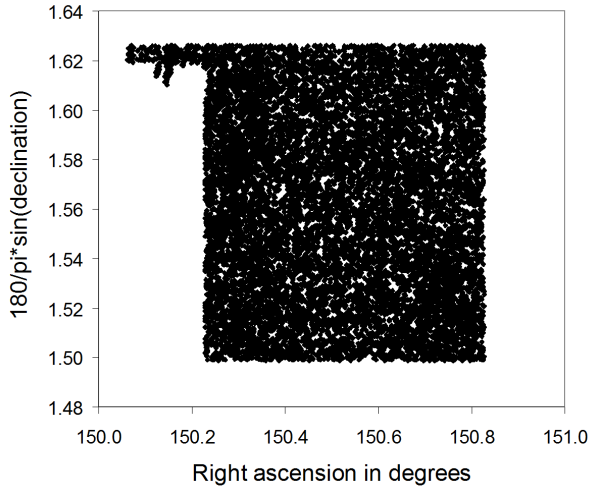


Fig. 1: Procedure to compute the spectroscopic area for the zCosmos survey as defined by the solid angle.

In the dopplerian redshift method, the relationship between redshifts and recession velocities is the following:

$$1 + z = \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}}. \quad (4)$$

From this equation, one may compute the recession velocity for a given redshift. Then the distance is computed as following:

$$distance = \frac{v}{H_o}. \quad (5)$$

From subsequent calculations an Ω_M of 0.19 was obtained which was used to derive the galactic density profile. Both methods give comparable distances with differences less than 5 % for redshifts up to 5.2 using Ω_M = 0.19. The difference between dopplerian and cosmological redshifts is discussed by Bedran [2]. Historically, the first solution to compute distances from cosmological redshifts was obtained by Mattig [9] which is based on Friedmann equations of general relativity. Mattig equation with q_o = 0.5 also provides distances close to what is obtained using dopplerian redshifts; however, Mattig had to assume that conservation of mass is applicable to the Universe in his derivations which is a big bang cosmology. On the other hand, dopplerian redshifts do not require any assumption on the cosmology, and present the advantage that they also explain blueshifts that are being observed such as for Andromeda.

2.1.3 Sources of data

The zCosmos galactic survey Data Release DR1 was used [8].

2.2 Propagation of light in an expanding space

The main hypothesis for the development of a model for the propagation of light in an expanding space, is that the speed of light is frame-independent. Considering redshifts, this means that the relative movement of a light source does not change the speed of light emitted; however, it does add or subtract energy to the photon. In a dopplerian world, this change in energy level changes the frequency of the source of light, and not the speed. However, as space between the photon and the observer expands, this expansion is added to the overall distance the photon has to travel in order to reach the observer - in over words the speed of light is frame-independent with respect to the local space. This implies that there exists a distance for which the recession speed between the observer and the photon equals the speed of light, which is the Hubble sphere, and that recession speed can exceed than the speed of light for large distances. The frame-independent hypothesis for the speed of light has been established in the past with the experiment of Michelson-Morley [10]. Based on observations of double stars [14, 4] it was shown that the velocity of propagation of light does not depend on the velocity of motion of the body emitting the light.

As a consequence of the above, LTDs are not anymore additive, meaning that if we have three points aligned in space, the distance between the two extremes is not anymore equal to the sum of the two sub-segments as measured in LTDs.

Based on the above hypothesis, the Euclidean distance between the photon and the observer is described by the following differential equation:

$$\frac{dy}{dt} = -c + H_o \cdot c \cdot T, \quad (6)$$

where y is the Euclidean distance between the photon and the observer, T the LTD between the observer and the photon, c the celerity of light, and H_o the Hubble constant.

2.3 Conversion of light travel distances to Euclidean distances

Let us consider a photon initially situated at a Euclidean distance y_o from the observer and moving at celerity c in the direction of the observer. Let us say T is the initial LTD between the photon and the observer, and define the Hubble constant function of LTDs.

The differential equation describing the motion of the photon in the LTD framework is described by eq. (6). By taking a reference point in time in the past, and T_b be today time from this reference point, we get $T = T_b - t$. Hence, $dt = -dT$. Therefore, eq. (6) becomes:

$$\frac{dy}{dT} = c - Ho \cdot c \cdot T, \quad (7)$$

with boundary conditions $y(T) = y_o$ and $y(0) = 0$.

By integration from 0 to T, the following relationship relating Euclidean distances y to light travel distances T is obtained:

$$y = c \cdot T - \frac{c \cdot H_o \cdot T^2}{2}. \quad (8)$$

The corresponding horizon computed by setting $\frac{dy}{dT} = 0$ is $T_h = \frac{1}{H_o}$ which is the Hubble sphere.

2.4 The Hubble constant was determined with respect to LTDs

In general the literature refers to the Hubble constant measured with respect to LTDs. A common way to obtain the Hubble constant is based on standard candles with supernovae and cepheids [13, 1] and the Tully-Fisher relation [5]. Both the standard candle and Tully-Fisher method rely on the distance modulus. As shown below the distance modulus gives a measure of LTDs and not Euclidean distances.

Let us recall the derivation of the distance modulus. The magnitude as defined by [12] is:

$$m = -2.5 \log F + K, \quad (9)$$

where m is the magnitude, F the brightness or flux and K a constant. The absolute magnitude is defined as the apparent magnitude measured at 10 parsecs from the source.

Planck's law for the energy of the photon leads to a redshift correction to the distance modulus

$$E = \frac{h \cdot c}{\lambda}, \quad (10)$$

where E is the energy of the photon, h the Planck's constant, and λ the light wavelength.

The ratio of observed to emitted energy flux is derived from eq. (10), leading to

$$\frac{E_{obs}}{E_{emit}} = \frac{\lambda_{emit}}{\lambda_{obs}} = \frac{1}{1+z}. \quad (11)$$

From geometrical considerations, the projected surface of the source of light on the receptor diminishes with a relationship proportional to the inverse of square distance from the source of light; hence, the following relationship is obtained for the brightness or flux:

$$F_{obs} \propto \frac{L_{emit}}{d^2} \cdot \frac{E_{obs}}{E_{emit}}, \quad (12)$$

where L_{emit} is the emitted luminosity and d the distance to the source of light.

Combining eq. (9), (11) and (12), we obtain:

$$m = -2.5 \log \left(\frac{L_{emit}}{d^2 \cdot (1+z)} \right) + K. \quad (13)$$

And, because z is close to zero at 10 Parsec:

$$M = -2.5 \log \left(\frac{L_{emit}}{100} \right) + K, \quad (14)$$

where M is the absolute magnitude.

Hence, the distance modulus, eq. (13) minus (14) is:

$$m - M = -5 + 5 \log d + 2.5 \log(1+z), \quad (15)$$

with d in parsec and log means the logarithm to base 10.

The expansion of the Universe adds up to the Euclidean distance, and therefore the apparent magnitude of the source of light is fainter than if no expansion was present.

2.5 Evolution of the galactic density assuming no new galaxy formation

Assuming cosmological redshifts we have:

$$1 + z = \frac{a_o}{a_1}, \quad (16)$$

where a_o and a_1 are respectively the present scale factor and the scale factor at z.

From the conservation of mass the density is proportional to the inverse of the cubic scale factor:

$$\rho \propto \frac{1}{a^3}. \quad (17)$$

Therefore, the model for the evolution of the density with respect to the present density is the following:

$$\rho_t = \rho_o \cdot (1+z)^3, \quad (18)$$

where ρ_t is the density in the past at redshift z and ρ_o is the present density.

3 Results

3.1 A flat density profile using Euclidean distances

Galactic density profiles have been derived for the two antagonistic scenarios respectively assuming that LTDs are additive, and with the propagation of light in an expanding space (figure 2). Note that the galactic density profiles obtained with cosmological redshifts and dopplerian redshifts are very similar. The highest redshift galaxies observed for the survey (z = 5.2) are very close to the Hubble sphere (which are at 13.65 Glyr) as calculated from cosmological redshifts with Ω_m=0.19.

The theoretical evolution of the galactic density with respect to the present density assuming no new formation of galaxies (figure 3) was computed assuming cosmological redshifts with eq. (18). Note that the first point in the galactic density profile is not representative of the average density as the sample volume is very small; hence, the measure represents the density in the neighbouring galactic cluster of the Milky Way (figure 2 and 3).

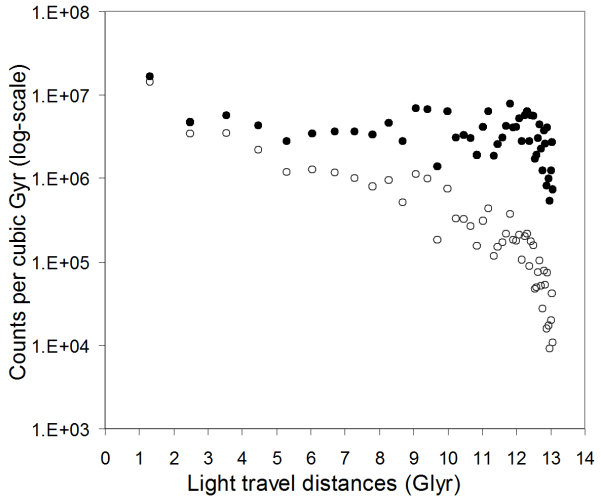


Fig. 2: Galactic density profile derived from the equivalent spherical sampling, where Glyr are billion light years from today. LTDs are obtained from redshift conversion with dopplerian redshifts. The blank dots indicate densities based on LTDs. The solid dots indicate densities obtained with Euclidean distances on the basis of dopplerian redshifts.

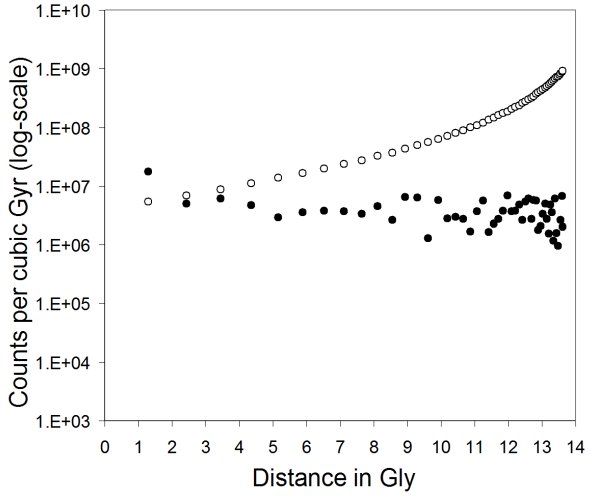


Fig. 3: Galactic density profile derived from the equivalent spherical sampling, where Gly are billion light years from today. LTDs are obtained from redshift conversion with cosmological redshifts (omega matter of 0.19).The solid dots indicate densities obtained with Euclidean distances on the basis of cosmological redshifts. The blank dots indicate the theoretical evolution of galaxies assuming that the survey is incomplete (with no new galaxy formation).

3.2 Estimation of Ω matter from galactic counts

The average galactic mass estimated from light deflection [15] is $1.7 \times 10^{11}\ M_{\odot}$. The Universe mean density is obtained by multiplying this figure with the average galactic count per cubic Glyr. Using dopplerian redshits the galactic count density is 4.6×10^{6} counts per cubic Glyr, leading to a mean Universe density of $1.84 \times 10^{-30}\ g/cm^3$. Using a Hubble constant of 71 $km/s/Mpc$ and recent estimates of the gravitational constant of $6.67 \times 10^{-8}\ cm^3/g/sec^2$ [11], the critical density is estimated at $9.47 \times 10^{-30}\ g/cm^3$ (from $\rho_c = \frac{3H^2}{8\pi G}$). Therefore, the corresponding Ω_m equals to 0.19. Note that smaller values of the Hubble constant would lead to a higher Ω_m.

3.3 Estimation of the number of galaxies in the visible Universe

Another challenge is to estimate the number of galaxies in the visible Universe. Using the galactic density in the nearby Universe from figure 2 expressed per cubic Glyr LTD, and the volume of the sphere of radius 14 Gly LTD, the number of galaxies in the visible Universe is estimated at 175 billion. Gott et al. [6] estimated a number of galaxies in the visible Universe at about 170 billion based on the Sloan Digital Sky Survey luminosity function data using the Press-Schechter theory. Both figures are consistent with each other; however, the author believes that these figures need to be reviewed to account only for the Euclidean radius when computing the volume of the visible Universe. As the galactic density profile is flat, it is expected that the estimated number of galaxies in the visible Universe is internally consistant with the bulk amount of galaxies observed in the survey converted to spherical values, i.e. multiplying the number of galaxies in the survey (10046 galaxies) by the sphere to survey solid angle ratio, which leads to 5.5 billion galaxies (see Table 1).

4 Discussion

A new approach is proposed in the present study to derive the galactic density profile which is based on the conversion of light travel distances to Euclidean distances. The method has been tested by computing the galactic density profiles based on the data from the zCosmos deep field survey.

In the scenario using LTDs with the sampling method, the galactic count per cubic Glyr grows according to a steep slope (figure 2), without accounting for the effect of the expansion which should add up to this growth. There is no explanation for such result - this scenario appears to be unrealistic. The scenario using Euclidean distances, shows a flat profile for the galactic counts per cubic Gyr (figure 3). However, there is still a gap between the computed galactic density profile and the theoretical evolution of galactic densities assuming no new galaxy formation. Leaving aside model bias, this gap may be interpreted as if galaxies grow in number over time. Another hypothesis is that the galactic survey is incomplete meaning that faint galaxies are left asside from the zCosmos survey at large distances, which would account for the missing galaxies causing the gap in figure 3. The theoretical density obtained by conservation of mass is too large by a factor

Table 1: Estimation of the number of galaxies in the visible Universe (radius 14 Glyr) using LTD distances and Euclidean distances.

	Radius of the visible Universe	Galactic density	Estimated number of galaxies
Using LTDs	14 Glyr	1.52×10^7 counts per cubic Glyr	175 billion
Using Euclidean distances with dopplerian redshifts	6.90 Glyr	4.60×10^6 counts per cubic Glyr	6.3 billion
Galaxy count of the survey converted to spherical values			5.5 billion

of order 200 at redshift 5.2. This discrepancy is unrealistically to large. Clearly more detailed work needs to be carried out to investigate this gap.

By applying conservation of mass, as we approach the singularity of the big bang, the Universe would have been so dense that it is difficult to explain how gravity did not prevent the early Universe from collapsing. A possibility is that the Hubble constant was much higher in the past leading to a higher critical density - cosmic inflation would still be necessary to overcome this issue. From the present study, the galactic density appears to be constant over time, which would corroborate the steady state cosmology of [3, 7]. The other condition being that the Hubble constant remains unchanged over time.

Acknowledgments

This work is based on zCOSMOS observations carried out using the Very Large Telescope at the ESO Paranal Observatory under Programme ID: LP175.A-0839.

Submitted on August 24, 2011 / Accepted on September 9, 2011

References

1. Altavilla G., Forentino, G., Marconi M., Musella I., Cappellaro E., Barbon R., Benetti S., Pastorello A., Riello M.,Turatto M., Zampieri L. Cepheid calibration of type Ia supernovae and the Hubble constant. *Monthly Notices of the Royal Astronomical Society*, 1982, v. 349 (4), 1344–1352.
2. Bedran M.L. A comparison between the Doppler and cosmological redshifts. *American Journal of Physics*, 2002, v. 70 (4), 406.
3. Bondi H., Gold T. The Steady-State Theory of the Expanding Universe. *Monthly Notices of the Royal Astronomical Society*, 1948, v. 108 (3), 252.
4. Brecher K. Is the Speed of Light Independent of the Velocity of the Source? *Physical Review Letters*, 1977, v. 39 (17), 1051–1054.
5. Freedman W.L., Madore B.F., Gibson B.K., Ferrarese L., Kelson D.D., Sakai S., Mould J.R., Kennicut Jr. R.C., Ford H.C., Graham J.K., Huchra J.P., Hugues S.M.G., Illingworth G.D., Macri L.M., Stetson P.B. Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant. *The Astrophysical Journal*, 2001, v. 553 (1), 47–72.
6. Gott R., Juric M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J. A Map of the Universe. *The Astrophysical Journal*, 2005, v. 624 (2), 463–484.
7. Hoyle F., Burbidge G., Narlikar, J.V. A quasi-steady state cosmological model with creation of matter. *The Astrophysical Journal*, 1993, v. 410, 437–457.
8. Lilly S.J., Le Fevre O., Renzini A., Zamoranti G., Scodeggio M., Contini T., Carollo C.M., Hasinger G., Kneib J.-P., Iovino A., Le Brun V., Maier C., Mainieri V., Mignoli M., Silverman J., Tasca L.A.M., Bolzonella M., Bongiorno A., Bottini D., Capak P., Caputi K., Cimatti A., Cucciati O., Daddi E., Feldmann R., Franzetti P., Garilli B., Guzzo L., Ilbert O., Kampczyk P., Kovac K., Lamareille F., Leauthaud A., Le Borgne J.-F., McCracken H.J., Marinoni C., Pello R., Ricciardelli E., Scarlata C., Vergani D., Sanders D.B., Schinnerer E., Scoville N., Taniguchi Y., Arnouts S., Aussel H., Bardelli S., Brusa M., Cappi A., Ciliegi P., Finoguenov A., Foucaud S., Franceschini R., Halliday C., Impey C., Knobel C., Koekemoer A., Kurk J., Maccagni D., Maddox S., Marano B., Marconi G., Meneux B., Mobasher B., Moreau C., Peacock J.A., Porciani C., Pozzetti L., Scaramella R., Schiminovich D., Shopbell P., Smail I., Thompson D., Tresse L., Vettolani G., Zanichelli A., Zucca E. zCosmos: A Large VLT/VIMOS Redshift Survey Covering 0 ¡ z ¡ 3 In The Cosmos Field. *The Astrophysical Journal Supplement Series*, 2007, v. 172, 70–85.
9. Mattig W. Über den Zusammenhang zwischen Rotverschiebung und scheinbarer Helligkeit. *Astronomische Nachrichten*, 1958, v. 284, 109.
10. Michelson A.A., Lorentz H.A., Miller D.C., Kennedy R.J., Hedrick E.R., Epstein P.S. Conference on the Michelson-Morley Experiment Held at Mount Wilson. *Astrophysical Journal*, 1927, v. 68, 341.
11. Mohr P.J., Taylor B.N. CODATA recommended values of the fundamental physical constants: 2002. *Reviews of Modern Physics*, 2005, v. 77, 1–107.
12. Pogson N. Magnitudes of Thirty-six of the Minor Planets for the first day of each month of the year 1857. *Monthly Notices of the Royal Astronomical Society*, 1857, v. 17, 12–15.
13. Schaefer B.E. The Peak Brightness of Supernovae in the U Band and the Hubble Constant. *The Astrophysical Journal*, 1995, v. 450, L5–L9.
14. De Sitter W. A proof of the constancy of the velocity of light. *Proceedings of the Section of Sciences - Koninkijke Academie van Wetenschappen te Amsterdam*, 1913, v. 15, 1297.
15. Tyson J.A., Valdes F., Jarvis J.F., Mills A.P.Jr. Galaxy mass distribution from gravitational light deflection. *The Astrophysical Journal*, 1984, v. 281, L59–L62.
16. Wright E.L. A Cosmology Calculator for the World Wide Web. *The Publications of the Astronomical Society of the Pacific*, 2006, v. 118 (850), 1711–1715.

An Analysis of States in the Phase Space: Uncertainty, Entropy and Diffusion

Sebastiano Tosto
E-mail: stosto@inwind.it

The paper aims to show the physical link between Fick's laws and entropy increase in an isolated diffusion system, initially inhomogeneous and out of the thermodynamic equilibrium, within which transport of matter is allowed to occur. Both the concentration gradient law and the entropic terms characterizing the diffusion process are inferred from the uncertainty equations of statistical quantum mechanics. The approach is very general and holds for diffusion systems in solid, liquid and gas phases.

1 Introduction

Diffusion concerns the transport of matter activated by thermal motion of atoms and molecules. Theoretical and experimental reviews on the mechanisms of mass transfer in solid, liquid and gas phases are widely reported in literature, e.g. [1, 2]. The importance of diffusion is well recognized in the kinetics of microstructural changes, nucleation of new phases, phase transformations, homogeneization and recrystallization of alloys and so on [3]; for instance electric conduction includes phenomena closely related to the transport mechanisms of ions and electrons.The theoretical background of the diffusion is based on an intuitive hypothesis: the driving energy that governs the mass transfer is related to the concentration gradient of molecules or atoms or ions in a diffusion medium, which can be simply the vacuum or a gas/liquid/solid phase. Such an assumption is so simple and reasonable to skip a more profound consideration just about the physical meaning of its general character. It is sensible to expect that this generality, and that of the related concentration gradient driving force itself, should be in fact consequence of some general principle of nature. This consideration recalls in effect the second law of thermodynamics, as concerns in particular the probabilistic character of the entropy. Consider an arbitrary number of particles "a" diffusing within a medium "b"; whatever the former might be, e.g. ions, atoms, molecules and so on, in the following they will be shortly referred to as particles, whereas the system formed by "a" and "b" will be referred to as diffusion system. One expects that after a proper time range, the system attains the most probable configuration, i.e. a uniform distribution of "a" into "b" regardless of the particular initial configuration assumed in general in a non-equilibrium state. So a net mass flow was necessarily occurring before reaching this limit situation, after which it is no longer allowed to occur. The entropy seems to be the thermodynamic concept most closely related to describe the transient and final configurations. This means that: (i) the dimensionless entropy formula $-\sum_i w_i \log(w_i)$, where the index i numbers the thermodynamic states allowed to the diffusing particles, should be involved since the beginning into the concentration gradient formulation of any diffusion problem; (ii) this formula should reduce to the simpler Boltzmann form $-\log(w_{eq})$ when the equilibrium configuration is effectively attained; (iii) the mass flow $\mathbf{J}$ is by consequence different from zero only during the time step (i), whereas it reduces to zero at the asymptotic time step (ii). Our knowledge on the diffusion process is thus based on a phenomenological hypothesis, the concentration gradient law, and on a general principle of nature, the entropy. It would be significant to regard both concepts as a natural consequence of a unique and more general principle of nature, without the need of phenomenological assumptions. Of course a general approach to this problem cannot leave out the quantum aspect of any problem inherent the dynamics of particles on microscopic scale. Justifying from the quantum point of view the concentration gradient driven diffusion law would provide a sound physical basis to the general problem of mass transport, whereas the continuity equation, if applicable, would also appear itself as a corollary identified by well-defined physical requirements about the diffusion system. On the one side it is certainly significant to demonstrate by means of a unique general principle the quantum origin of the macroscopic equations describing how the configuration of the diffusion system evolves as a function of time because of the mass transfer. On the other side this task seems further noteworthy if carried out within the same theoretical frame that allows describing the quantum properties of matter. The purpose of the present paper is to investigate the quantum basis hidden into the gradient law, i.e. to demonstrate that the uncertainty is the basic quantum principle leading to the first Fick law as a corollary. Moreover the theoretical model proposed here also confirms through a simple and straightforward approach that the entropy of the diffusion system is the other key concept underlying the mechanisms of mass transport.

2 Classical background

For simplicity, let us regard the diffusion system as an isolated thermodynamic system formed by an isotropic body of matter and introduce the mass flow as follows:

$$\mathbf{J} = c\mathbf{v}, \tag{2,1}$$

where c is the concentration or more in general the activity of the diffusing particle and $\mathbf{v}$ its displacement velocity. Eq.

2,1 is simply a definition. A further equation appears necessary to introduce a physical hypothesis about the thermodynamic force **F** that triggers the flow. Expressing this hypothesis through the following equation, known as first Fick law

$$\mathbf{J} = -D\nabla c, \qquad c = c(x,t) \tag{2,2}$$

and combining these equations, one finds indeed

$$\mathbf{v} = -\frac{D}{k_B T}\nabla[k_B T \log(c/c_0)], \qquad c_0 = c_0(t), \tag{2,3}$$

where c_0 is an arbitrary reference concentration not dependent upon x but possibly dependent on time. The definition of mobility β of the diffusing particle

$$\mathbf{v} = \beta\mathbf{F} \tag{2,4}$$

entails therefore as a consequence at constant T

$$D = \beta k_B T, \qquad \mathbf{F} = -\nabla[k_B T \log(c/c_0)]. \tag{2,5}$$

One finds therefore through the definition of mobility both the sought force, which reasonably results equal to the gradient of the potential energy $\mu = k_B T \log(c/c_0)$, and the well known Einstein equation linking β to D. The form of **F** provides a partial answer to the aforesaid point (iii): if c is equal everywhere in the diffusion system, then it does not longer depend upon x; so, defining c_0 equal or proportional to this uniform limit value of c, one finds $\mathbf{F} = 0$ and thus $\mathbf{v} = 0$ everywhere. This shows that **F** accounts for the net mass flow in the diffusion system until $c \to c_0$. These preliminary considerations highlight that the diffusion law can be effectively related to a thermodynamic function, the chemical potential, that describes the driving force allowing the transport of matter. Exploit now again the basic definition eq. 2,1 to evidence how arbitrary changes of both c and **v** affect **J**. Consider then

$$\delta\mathbf{J} = \mathbf{v}\delta c + \delta\mathbf{J}', \qquad \delta\mathbf{J}' = c\delta\mathbf{v} \tag{2,6}$$

in the time range δt during which $\delta\mathbf{J}$ is allowed to occur. Note that δc can be due: (i) to the change δm of m within the reference volume V defining c or (ii) to the change δV of V for fixed m or (iii) to both reasons. In any case, defining the space range $\delta x = v_x\delta t$ where the particles are allowed to diffuse along the x-direction during δt, the x-component of eq. 2,6 reads $\delta J_x/\delta x = \delta c/\delta t + c\delta v_x/\delta x$. So, for infinitesimal changes dc and $d\mathbf{v}$ of the process parameters and of the dynamical variables dt and dx, the last equation reads $\nabla\cdot\mathbf{J} = \partial c/\partial t + c\nabla\cdot\mathbf{v}$, i.e. in general

$$\nabla\cdot\mathbf{J} = \frac{\partial(c+C)}{\partial t}, \qquad C = \int_{t_o}^{t} c'\nabla\cdot\mathbf{v}'dt', \qquad C = C(x,t) \tag{2,7}$$

with the integral calculated between the fixed time t_o, e.g. the beginning of the diffusion process, and the current time t. If holds the condition $\nabla\cdot\mathbf{v} = 0$, then $\nabla\cdot\mathbf{J} = \partial c/\partial t$ describes a particular diffusion process where the rate of concentration change is equal to the gradient of related mass flow, which necessarily means lack of sinks or sources of matter within the volume element where is defined c. Since $c\nabla\cdot\mathbf{v}$ results because of the term $\delta\mathbf{J}'$ additional to $\delta\mathbf{J}$, it appears that the well known second Fick equation is a particular case of eq. 2,6 for $\delta\mathbf{J}' = 0$. Actually $\delta\mathbf{J}' \neq 0$ is due not only to a possible chemical reaction that involves the diffusing particle and modifies the local concentration of the diffusion system but, more in general, also to any local force field that attracts or repels the diffusing particles and perturbs their motion. Note indeed that $\delta\mathbf{J}' = c\mathbf{a}\delta t = \mathbf{F}'_V\delta t$ yields

$$\frac{\delta\mathbf{J}'}{\delta t} = \frac{\mathbf{F}'}{V} = \mathbf{F}'_V,$$

being in general $\mathbf{F} \neq \mathbf{F}'$. The force per unit volume $\mathbf{F}'_V$ that controls the perturbation term $\delta\mathbf{J}'$, appearing in eq. 2,6 as a perturbation of **J** is particularly interesting for charged particles diffusing in an ionic medium where polarized impurities are active. Note indeed that $\mathbf{v}\cdot\mathbf{J}$ has physical dimensions of energy per unit volume; then $\mathbf{v}\cdot\delta\mathbf{J}' = (m\delta v^2/2)V^{-1}$, i.e. the effect of $\mathbf{F}'_V$ is that of perturbing the kinetic energy of the particle in the interaction volume V. It is usually acknowledged that the time enters into the diffusion equation thanks to the continuity condition that leads to the second Fick law. Yet the mere definition of eq. 2,1 entails an interesting conclusion: regardless of the aforesaid effects related to δm that possibly alter the plain diffusion process, the time evolution of the system is actually consequence of the concentration gradient law; although the Fick hypothesis does not contain explicit reference to the time, this latter enters indeed into the problem through **v**. The present considerations show therefore that the ancillary condition of continuity is not necessary to infer the second Fick law; rather, simply taking into account the finite range δt required to justify $\delta\mathbf{J}$, as nothing changes instantaneously in nature, the continuity condition appears to be itself a corollary of the definition of mass flow and not an additional boundary condition. Otherwise stated, even from a merely classical point of view the time coordinate appears a necessary ingredient together with the space displacement to account for the mass transfer in any diffusion problem; consequently the position $\nabla\cdot\mathbf{v} = 0$ does not represent a supplementary hypothesis "ad hoc" but simply a possible chance allowed for $\delta\mathbf{J}$. This conceptual basis, to be further implemented by quantum considerations reasons in the next section, is characterized by three physical features summarized as follows: (i) the definition of mass flow, eq. 2,1; (ii) the gradient concentration law; (iii) the necessity of introducing diffusion driven displacement $\delta\mathbf{r}$ and time range δt linked by $\delta\mathbf{r} = \mathbf{v}\delta t$, which also introduces the energy range $\delta\varepsilon = (\mathbf{v}\cdot\delta\mathbf{J})V$ corresponding to $\mathbf{F}\cdot\delta\mathbf{r}$ within the reference volume V defining c.

3 Preliminary quantum considerations

This section introduces the basic ideas to describe the diffusion system according to the uncertainty relationships

$$\Delta x \Delta p_x = n\hbar = \Delta t \Delta \varepsilon, \tag{3,1}$$

where n is an arbitrary number of quantum states allowed to any particle moving in the space range Δx with conjugate momentum falling in the momentum range Δp_x; the ranges are taken positive by definition. As already shown in [4], the second equality is obtained from the first one defining formally $\Delta t = \Delta x / v_x$ and $\Delta \varepsilon = \Delta p_x v_x$ linked by the same n; v_x is the velocity with which the particle travels within Δx. No hypothesis is required about the ranges that quantify the concepts of space and time uncertainty. Their sizes, in principle arbitrary, can vary from zero to infinity; moreover nothing is known about their analytical form, e.g. any local functional relationship like $p_x = p_x(x)$ within Δx is physically meaningless because both p_x and x are assumed random, unknown and unpredictable. Yet, despite such an agnostic point of view, relevant features of the ranges are apparent. First, v_x must be upper bounded. Consider a free particle in finite sized Δx and Δp_x with n finite as well: if $v_x \to \infty$ then $\Delta t \to 0$ would require $\Delta \varepsilon \to \infty$, which in turn would allow in principle an infinite energy ε; but this is impossible once having merged both uncertainties via a unique n, as $\varepsilon \to \infty$ is inconsistent with any p_x falling within the finite range Δp_x and thus necessarily finite itself. Hence the simple fact of having regarded together space and time uncertainties, i.e. admitting that both dynamical variable concur to describe any physical system, requires $v_x \leq v_x^{\max}$; eqs. 3,1 entail as a corollary the well acknowledged existence of an upper limit for the propagation rate of any signal. Moreover put $\Delta x = x - x_o$ and consider that the coordinate x_o, whatever it might be, is defined in an appropriate reference system that defines position and size of Δx and v_x as well; yet, being x_o indeterminate and indeterminable, the present approach based on Δx only does not specify in fact any particular reference system. The same holds for course also for the other ranges of eqs. 3,1, in particular for the time frame. Also, in lack of constrains or hypotheses the reference system could be in principle Cartesian or curvilinear or inertial or non-inertial or anything else. This means that any physical problem discarding "a priori" the local dynamical variables and exploiting eqs. 3,1 only, i.e. replacing

$$x \to \Delta x, \qquad p_x \to \Delta p_x, \qquad t \to \Delta t \tag{3,2}$$

holds by definition in any space-time reference system R. Hence eqs. 3,1 entail that all reference systems are indistinguishable and thus equivalent in describing the properties of quantum particles. If so, it eventually follows that the upper value allowed to v_x, whatever it might be, must be invariant in any R. Indeed v_x is defined by its own reference system; being the former arbitrary, the latter is arbitrary as well. Consider instead a well specified value of v_x, e.g. just its maximum value $v_x^{\max}$; this latter must be uniquely defined in R and in any other R' otherwise R and R' could be identified depending on their own $v_x^{\max}$, e.g. because of a greater velocity allowed in either of them, thus contradicting their indistinguishability. It appears therefore that equivalence of all reference systems and invariance of $v_x^{\max}$ are strictly linked. The time coordinate, previously introduced to account for the finite rate with which occurs the mass flow change $\delta \mathbf{J}$, still appears here as a consequence of the finite velocity v_x with which any particle moves within Δx and entails a finite time range to change the configuration of the diffusion system. Yet now Δt takes a more general physical meaning, as it appears from the previous considerations and it will be shown in the next sections. The uncertainty inherent eqs. 3,1 requires innately a time range for particles delocalized in Δx, i.e.: any physical process characterized by an energy spread $\Delta \varepsilon$ requires a time range Δt during which is to be expected a momentum change falling within Δp_x too. Previous papers [5, 6] have shown that this way of regarding eqs. 3,1 is enough to calculate the energy levels of hydrogenlike and many electron atoms/ions and diatomic molecules without solving any wave equation; then is attracting the idea that even the diffusion model can be formulated in terms of particles randomly spreading within their own delocalization space ranges conceptually arbitrary, unknown and unknowable themselves. As in the quoted papers, the statistical formulation of the quantum uncertainty is the only assumption necessary also in the context of the present problem. Suppose of having N particles in N_V elementary volumes Δx^3 of diffusion medium at a fixed time of the diffusion process. Regardless of the equilibrium or non-equilibrium situation at the given time, let

$$W_{cl} = \binom{N}{N_V}, \quad N = N(t), \quad N_V = N_V(t), \quad V = \Delta x^3 \tag{3,3}$$

be the number of ways to distribute N classical particles in N_V available sites of the diffusion medium. From a quantum point of view the combinatorial calculus still holds in principle also in the case of identical particles, as it is done in the Fermi-Dirac and Bose-Einstein statistics; one must simply replace W_{cl} with the pertinent expressions of numbers of states taking into account the indistinguishability of identical particles. Note in this respect the characteristic way of working of eqs. 3,1: once accepting the replacements 3,2, the physical interest about the system moves from the constituent particles to their phase space. On the one side just this feature of eqs. 3,1 entails the corollary of quantum indistinguishability of identical particles when considering uniquely ranges of dynamical variables where *any* particle could be found, rather than the actual dynamical variables of the particle itself; indeed this latter is never specified "a priori". On the other side this explains the general worth of the eqs. 3,1 regardless of the specific system concerned: the present model

holds in principle for diffusion processes in solid or liquid or gas phase, since no hypothesis is formulated about N and N_V of W_{FD} or W_{BE}. Further information on the process, e.g. the role of lattice defects on the effectiveness of mass transport, are to be introduced "a posteriori" through specific values of the coefficient D only, see eq. 2,2, whose quantum root will be indeed highlighted in the next section. It is important however that regardless of the kind of diffusion system, the computation of the number of allowed states accessible to the particles requires calculating the ways of distributing N objects into N_V volume elements of sizes $\Delta x^3_{1\leq i\leq N_V}$; this is possible even in the present approach because the combinatorial computation of allowed states does not require knowing where exactly are located these volumes in the diffusion system, which indeed would be prevented by eqs. 3,1. Just this computation yields the corresponding entropy of the diffusion system. At the very beginning of the diffusion process one can imagine an isolated ordered system S_0 where all particles are confined in some arbitrary volume of the system; as the particles are allowed to walkover randomly to occupy a greater volume, the number of allowed thermodynamic states progressively increases as a function of time. For $t \to \infty$ the system reaches an asymptotic state S_∞ to which corresponds a net mass flow $\mathbf{J} = 0$. The driving force of the diffusion process is thus certainly correlated to the tendency of the system towards its state of thermodynamic equilibrium and maximum entropy. Thus eqs. 3,3 simply tell that in non-equilibrium conditions the system $S(t)$ at the time t is such that $S_0 \leq S(t) < S_\infty$, until the distribution of particles corresponds to the maximum number of quantum states inherent $S_\infty \neq 0$; correspondingly $\mathbf{J} \neq 0$ describes net mass flow in the system tending the maximum entropy, until when $\mathbf{J} \to 0$. The next section aims to show that this intuitive picture of diffusion process will be inferred together with the concentration gradient law through eqs. 3,1 only, without need of any phenomenological hint.

4 Diffusion quantum model

By definition the uncertainty ranges of eqs. 3,1 include any position and momentum of the particles during the diffusion process, despite both dynamical variables are expected to change as a function of time by effect of an appropriate driving force $\mathbf{F}$. In principle one could think Δx and Δp_x large enough to include any possible change of x and p_x from the initial stage of the diffusion process to the final state of thermodynamic equilibrium; indeed the eqs. 3,1 admit possible interactions of these particles with the surrounding medium along the diffusion path $\delta\Delta x = v_x\delta t$ from $\delta t = 0$ to $\delta t \to \infty$, e.g. by elastic and anelastic collisions, through an appropriate size of the energy range $\Delta\varepsilon$. Owing to the complete arbitrariness of the ranges, however, this approach although sensible does not appear far reaching to get relevant information about the process. Yet it is also possible, and more heuristic, to require that Δx and Δp_x are allowed to change themselves as a function of time without contradicting their arbitrariness and without requiring any information on the local values x and p_x; in effect eqs. 3,1 can be differentiated with respect to t and x whatever the current time and space coordinates of particles might be. Consider thus $\delta\Delta x$ and $\delta\Delta p_x$, rather than δx and δp_x, regardless of whether the displacement of matter from two different points of the diffusing medium occurs with or without net mass flow; $\delta\Delta x$ describes the change of delocalization range to which is related the assumed change of momentum $\delta\Delta p_x$ by effect of $\mathbf{F}$. The force is here easily justified by eqs. 3,1 themselves, regardless of other specific motivations: $\Delta\dot{x}$ defining $\delta\Delta x = \Delta\dot{x}\delta t$ requires $\Delta\dot{p}_x$, which therefore affects the range of values allowed to any p_x; in turn the change of p_x, allowed to occur and thus in fact occurring, entails $F_x = m\partial v_x/\partial t$. Since it is possible to write $\delta\Delta p_x = (\partial\Delta p_x/\partial t)\delta t$, then

$$\frac{\partial\Delta p_x}{\partial t} = -n\hbar\Delta x^{-2}v_x = F_x = m\frac{\partial v_x}{\partial t}, \qquad v_x = \frac{\partial\Delta x}{\partial t}. \quad (4,1)$$

Note that here v_x is not the diffusion velocity of the particle but the rate with which changes Δx, so F_x is defined in the phase space of the particle. Yet this information is enough as concerns the diffusion problem: by effect of F_x the particle is allowed to move faster, being however still delocalized within the larger range $\Delta x' = \Delta x + \delta\Delta x$. This is why the momentum of the particle is allowed to change along with $\delta\Delta x$. The notation of velocity is unique to emphasize that v_x of eq. 4,1 and v_x of the particle defining eqs. 3,1 are both arbitrary and thus assumed coincident. On the one side this representation is consistent with well known ideas of the diffusion process, e.g. particle jumps through different sites in a crystal lattice or particle collisions randomly occurring in gas phase; on the other side it suggests that the local concentration change is described by a constant amount of mass m allowed to move slower or faster in a decreasing or increasing phase space delocalization range depending on the sign of the velocity component v_x. In this way the force component F_x introduced via the deformation of the momentum range is conceptually consistent with that of eq. 2,5: to the momentum change rate that defines the classical force corresponds now, from the point of view of eqs. 3,1, the existence of a force field $\Delta\dot{p}_x$ necessary to account for any possible $\dot{p}_x$ during the diffusion process. Let us differentiate now eqs. 3,1 with respect to x to link the change of size of the delocalization range $\delta\Delta x$ and that of the momentum range $\delta\Delta p_x$ when the particle displaces by δx; this yields

$$\frac{\partial\Delta p_x}{\partial x} = -n\hbar\Delta x^{-2}\frac{\partial\Delta x}{\partial x}. \quad (4,2)$$

Eqs. 4,1 and 4,2 describe the dynamics of the diffusing particle as a function of time in agreement with eqs. 3,1. The classical eqs. 2,6 and 2,7 have introduced $\mathbf{v}$ as macroscopic average velocity describing the net mass flow due to

the displacement rate of the particle; now the quantum approach shows how the uncertainty compels regarding a random mass flow in the phase space of the particle: the deterministic force of eq. 2,5, exactly defined at any point of the diffusion system, is now replaced by the random force of eq. 4,1 controlled by arbitrary values of n and Δx. Let us show now that this agnostic point of view, far from being elusive of the problem, is actually source of relevant physical information. The fact that the diffusion is allowed in a given volume $V = (n\hbar)^3 \Delta p_x^{-3}$ suggests exploiting an approach conceptually identical but formally different from that introduced in section 2. If the motion of the particle is random, the orientation of its momentum $\mathbf{p}$ is defined in general within a sphere of radius $|\Delta\mathbf{p}|$ whose volume is thus $\propto \Delta p_x^3$ once taking $\Delta p_x \equiv |\Delta\mathbf{p}|$; since the medium is isotropic and the uncertainty ranges are arbitrary and unknown, there is no necessity to introduce explicitly separate ranges Δp_x, Δp_y and Δp_z. So, instead of starting from $\partial \Delta p_x / \partial x$, it is more convenient considering $a''' \Delta p_x^2 \partial \Delta p_x / \partial x$, where a''' is a proper proportionality factor; indeed $\hbar^{-3} \Delta p_x^2 d\Delta p_x$ is proportional to the number of particles whose momentum was initially included in a sphere of radius Δp_x and takes after the time range δt values falling in the section of sphere between Δp_x and $\Delta p_x + d\Delta p_x$. So introducing the quantity $a'' \partial \Delta p_x^3 / \partial x$ means considering a volume element in the momentum space of the particle, which yields in turn with the help of the eq. 3,1 $a' \partial \Delta x^{-3} / \partial x$; here a'' and a' are trivial numerical factors. In conclusion, although starting from a 1D equation, we have introduced a volume element $V = \Delta x^3$ that represents an elementary volume of the diffusion medium where is located a given amount of diffusing mass m corresponding to the concentration c. This defines the equation

$$-\frac{a'}{V^2}\frac{\partial V}{\partial x} = \frac{a'}{V}\frac{\partial \log(V_o/V)}{\partial x}, \qquad (4,3)$$
$$V = V(x,t), \qquad V_o = V_o(t),$$

where the arbitrary constant V_o is a reference volume by definition not dependent on x but possibly dependent on t. Consider first the left hand side of this identity, which reads

$$-\frac{a'}{V^2}\frac{\partial V}{\partial x} = -\frac{a' m}{V^2}\frac{\partial c^{-1}}{\partial x} = \frac{a' m}{c^2 V^2}\frac{\partial c}{\partial x} = \frac{a'}{m}\frac{\partial c}{\partial x},$$
$$c = \frac{m}{V}, \qquad c = c(x,t),$$

where c has here the same physical meaning introduced in the early eq. 2,1, although the equation concerns now the phase space rather than a selected volume of matter. This result regards m as a constant with respect to x, i.e. c depends on x through the volume Δx^3 around m only. This point of view, extended to various volumes Δx_i^3 in which the diffusion medium can be ideally divided, entails that the deformation extents $(\Delta x_i + \delta \Delta x_i)^3$ change as a function of x in order that the respective δc_i represent by consequence these changes; this holds when a total amount of matter $\sum_i m_i$ is simply redistributed along x, thus changing the reference volumes that physically define the respective c_i only, or when $\sum_i m_i$ is subjected to change itself because of sinks or sources of matter in the diffusion medium; this is why the time has been explicitly introduced in eqs. 3,3. The right hand side of the first eq. 4,3 depends certainly upon time through V_o; the same holds therefore for the left hand side, i.e. $a' = a'(t)$. Moreover a' depends in general on x as well; indeed it accounts for how $\partial \Delta x^{-3} / \partial x$ changes in general as a function of x, so $a' = a'(x,t)$. Eventually a' must be consistent with the idea of a mass m crossing the momentum space surface proportional to Δp_x^2 during the time range δt, i.e. the physical dimensions of a' must be $\mathrm{mp^2t} = \mathrm{ml^2t^{-1}}$ like that of $\hbar$; this point will be better emphasized in section 5. Specifying thus purposely the proportionality factor a' in order that also the right hand side of eq. 4,3 depends on c, one finds

$$J_x = -D\frac{\partial c}{\partial x}, \qquad a' = -Dm, \qquad D = D(x,t). \qquad (4,4)$$

The physical dimensions of D are therefore $\mathrm{l^2t^{-1}}$. This result represents the first task of the present paper: to infer the concentration gradient law governing any diffusion process as a consequence of the fundamental eq. 3,1, thus showing the quantum origin of the first Fick law. To proceed further, consider now the right hand side of eq. 4,3 rewritten with the help of the second eq. 4,4 as

$$J_x = -Dc_o f \frac{\partial \log(f)}{\partial x}, \qquad f = \frac{c}{c_o}, \qquad c_o = \frac{m}{V_o}, \qquad c_o = c_o(t).$$

The first expression calculated in an arbitrary point $x = x_a$ defines $f = f_a$ through the local concentration c_a and reads, with obvious meaning of symbols,

$$J_a = -D_a c_o f_a \left.\frac{\partial \log(f)}{\partial x}\right|_{f_a} = -D_a \left.\frac{\partial c}{\partial x}\right|_{x=x_a}, \qquad (4,5)$$
$$f_a = \frac{c_a}{c_o}, \qquad D_a = D(x_a,t).$$

Let us expand in series the function $\log(f)$ around x_a

$$\log(f) = \log(f_a) +$$
$$+ \left.\frac{\partial \log(f)}{\partial x}\right|_{f_a}(x - x_a) + \frac{1}{2}\left.\frac{\partial^2 \log(f)}{\partial x^2}\right|_{f_a}(x - x_a)^2 + \ldots$$

and calculate this expression in another point x_b, arbitrary as well; this yields

$$\left.\frac{\partial \log(f)}{\partial x}\right|_{f_a} = \frac{\log(f_b) - \log(f_a)}{x_b - x_a} - \frac{1}{2}\left.\frac{\partial^2 \log(f)}{\partial x^2}\right|_{f_a}(x_b - x_a) - \ldots,$$
$$f_b = \frac{c_b}{c_o}.$$

Replacing in eq. 4,5 and putting $J_o = -D_a c_o/(x_b - x_a)$ one finds

$$\frac{J_a}{J_o} = -f_a \log(f_a) + \\ + \left(f_a \log(f_b) - \frac{f_a (x_b - x_a)^2}{2} \left. \frac{\partial^2 \log(f)}{\partial x^2} \right|_{f_a} + \cdots \right). \tag{4,7}$$

Rewrite now c_o not yet defined as $c_o = (c_b - c_a)/\gamma$, being γ a dimensionless proportionality factor; this position entails

$$J_o = -\frac{D_a}{\gamma} \frac{c_b - c_a}{x_b - x_a}, \tag{4,8}$$

$$f_a = \gamma \frac{c_a}{c_b - c_a}, \qquad f_b = \gamma \frac{c_b}{c_b - c_a}, \qquad \gamma = \gamma(t).$$

The last position agrees with the dependence of c_o upon time through V_o. In this way J_o agrees conceptually with J_a and thus with the definition of concentration gradient driven mass flow yet with a different diffusion coefficient $D_o = \gamma^{-1} D_a$; it reduces indeed to the usual differential form $J_o = -D_o \partial c/\partial x$ in the limit $x_b \to x_a$ that necessarily entails $c_b \to c_a$. One would expect that in this limit $J_o \to J_a$, which should require $\gamma \to 1$; however the fact that in general $\gamma \neq 1$, as it is shown below, suggests that J_o is physically consistent with but numerically different from J_a. Before concerning this point, note that the second and third eqs. 4,8 require $f_b = \gamma + f_a$; so eq. 4,7 reads

$$\frac{J_a}{J_o} = -f_a \log(f_a) + \left(f_a \log(f_a + \gamma) - d_{ab}^2 \left. \frac{\partial^2 \log(f)}{\partial x^2} \right|_{f_a} \right),$$

$$d_{ab}^2 = f_a \frac{(x_b - x_a)^2}{2}, \tag{4,9}$$

having neglected for simplicity the higher order terms of series development of $\log(f)$. The time function γ is therefore a parameter controlling the evolution of the ratio J_a/J_o, which results to be also a function of $x_a - x_b$ and $c_a - c_b$ via f_a. To explain this result, let x_b be the coordinate of a particle at the beginning of the diffusion process and x_a that of the particle at a later time, while c_b and c_a are the respective concentrations. In general $f_a \neq f_b$ for $x_a \neq x_b$ since $c_a \neq c_b$. Consider however in this respect the particular limit condition $c_b \to c_a$ to be expected in two relevant cases: (i) at the very beginning of the diffusion process, when the particle has traveled an infinitesimal path so that x_a is very close to its initial position x_b; (ii) at the end of the diffusion process, when the particle has traveled a finite path with x_a arbitrarily far from x_b but the concentration is uniform throughout the diffusion system. In both cases it is convenient to define $\gamma \to 0$ in order that the undetermined form $\gamma/(c_b - c_a) \to 0/0$ does not necessarily cause divergent values of f_a and f_b. If $c_b \to c_a$ simply because $x_b \to x_a$, case (i), elementary manipulations of eq. 4,9 show that both sides tend to γ provided that $\gamma/f_a \to 0$; in effect this is verified because by definition $\gamma/f_a = (c_b - c_a)/c_a$, see eq. 4,8. The result is thus

$$\lim_{\substack{c_b \to c_a \\ x_b \to x_a}} \frac{J_a}{J_o} = \gamma, \qquad t \to 0, \qquad \gamma \to 0. \tag{4,10}$$

This simply means that at $t = 0$ there is no net flow of matter as $J_a = 0$. This is reasonable, because after a very short path the particle has high probability to return to its initial position. The second chance for $c_b \to c_a$ even though $x_a \neq x_b$ yields, putting again $\gamma \to 0$,

$$\lim_{\substack{c_b \to c_a \\ x_b \neq x_a}} \frac{J_a}{J_o^{eq}} = \frac{D_a}{J_o^{eq}} \left. \frac{\partial c}{\partial x} \right|_{x = x_a} = \gamma - d_{ab}^2 \left. \frac{\partial^2 \log(f)}{\partial x^2} \right|_{f^{eq}}, \tag{4,11}$$

$$t \to \infty, \qquad \gamma \to 0.$$

Note that γ can fulfill both conditions if its form is, for instance, like $t/(t^2 + t_o)$. Also note that in fact the behavior of γ can be consistent with any $c_b - c_a$, i.e. whatever this limit might be depending on the kind of diffusion system; being γ defined here by its limit condition only, one could hypothesize any stronger/weaker time dependence, e.g. like $t^k/(t^{k+1} + t_o)$, with k ensuring a finite value of $\gamma(c_b - c_a)^{-1}$ no matter how rapidly $c_a \to c_b$ case by case. Put therefore by definition

$$\lim_{c_a \to c_b} \gamma/(c_b - c_a) = \gamma_{ab}, \qquad \gamma_{ab} \neq 0. \tag{4,12}$$

The left hand side of eq. 4,11 has now the form

$$(x_b - x_a)\gamma(c_b - c_a)^{-1} (\partial c/\partial x)_{x = x_a}.$$

The right hand side vanishes for $\gamma \to 0$ if $c_a = c_b = const$ everywhere in the diffusion system because f is now a constant defined by the limit $f_a \to f^{eq}$, whence the notation J_o^{eq}. Hence $x_a \neq x_b$ and γ such that γ_{ab} remains finite require $\partial c/\partial x$ vanishing at x_a. As expected, the situation of uniform concentration entails on microscopic scale the asymptotic condition of thermodynamic equilibrium without net mass transfer. Hence the maximum chance of displacement is expected at times intermediate between 0 and infinity. If c_a is the same everywhere because x_a is arbitrary, then actually neither side of eq. 4,11 depends on x; so must hold also on a macroscopic statistical scale the conclusion that a uniform distribution of particles in the diffusion system makes the ratio J_a/J_o^{eq} of eq. 4,9 inconsistent with a net flow of particles. In fact this requires verifying that also the sum of all terms of eq. 4,9 over the indexes a and b fulfills the condition

$$\sum_{b,a} \lim_{\substack{c_b \to c_a \\ x_b \neq x_a}} \frac{J_a}{J_o^{eq}} = 0, \qquad t = \infty, \tag{4,13}$$

whereas in general, since f_a never diverges,

$$\sum_{b,a} \frac{J_a}{J_o} \neq 0, \qquad t > 0. \tag{4,14}$$

Actually the sums are extended to all paths of particles from the respective starting points x_b to their end points x_a, which also means summing over all elementary volumes $V_a = \Delta x_a^3$ and $V_b = \Delta x_b^3$ of the diffusing medium in which the particles are found with corresponding concentrations c_a and c_b; since both coordinates are arbitrary, this picture represents in fact any path between any points in the diffusion system. Before demonstrating eq. 4,13, note that the sum has conceptual meaning because in fact it does not require computing anything; it is introduced in principle because neither x_a nor x_b are known but are merely referred to their own V_a and V_b only, wherever their position in the diffusion system might be. Also note that the ratio J_a/J_o entails two harmonized but different definitions of mass flow: at numerator appears a local term, characterized by a concentration difference between two coordinates infinitely close each other, at denominator a macroscopic term characterized by coordinates arbitrarily apart. The flow described by J_a is thus a net flow of matter only controlled by D_a, since by definition an effective concentration gradient corresponds to it. The fact that the sum of ratios is finite in eq. 4,14 and equal to 0 in eq. 4,13 suggests that J_o must concern a macroscopic diffusion term controlled by $D_o = D_a\gamma^{-1}$, describing total displacement of matter that consists in principle of both vanishing and non-vanishing net mass flows because $J_o \neq 0$ even though $J_a = 0$; both flows are in fact allowed to occur in a macroscopic volume of diffusion system, so that neither of them can be excluded. Hence the ratio J_a/J_o in eq. 2,3 represents a sort of "displacement efficiency" corresponding to the thermodynamic force F_x of eq. 4,1, i.e. the chance that the random motion of particles produces an effective flow of matter between two arbitrary volumes within the diffusion system. Eq. 4,13 is then easily justified noting that J_o^{eq} changes sign by exchanging x_a and x_b if $c_a = c_b$, whereas J_a does not for the simple reason that its definition has nothing to do with x_b. In effect just the presence of a concentration gradient makes the environment around the coordinates x_a and x_b physically different; if the coordinates belong to different volumes V_a and V_b that define the respective non-equilibrium concentrations, the displacement of a particle between two points out of the equilibrium is distinguishable from that obtained keeping fixed c_a and c_b with reversed path. Instead the sums $\sum_{a,b}$ and $\sum_{b,a}$ at the equilibrium must be in principle identical, because a uniform distribution of particles within the diffusion system makes indistinguishable starting points and end points; if the diffusion system is perfectly homogeneous, then all volumes $V_i = \Delta x_i^3$ where $c \neq 0$ are identical. This is consequence of having defined c as due to a unique value of m into different volumes of phase space that define V_a and V_b of the diffusing medium. Thus the only chance for a sum to coincide with its own value of opposite sign is that the sum is null. Eq. 4,13 is in fact possible from a mathematical point of view because

$$\partial^2 \log(f)/\partial x^2 = -f^{-2}(\partial f/\partial x)^2 + f^{-1}\partial^2 f/\partial x^2, \qquad (4,15)$$

i.e. the former addend is certainly negative whereas the second can take in principle both signs; hence in principle the sum of terms at right hand side of eq. 4,11 can vanish for an appropriate value of $f_a = f_b = f^{eq}$. Let us return now to eq. 4,9 and note with the help of eq. 4,8 that for $f_a = 0$, i.e. $c_a = 0$, the ratio J_a/J_o is identically null in agreement with its probabilistic meaning. Then, since each coordinate x_a belongs to its own volume V_a that defines c_a, summing over all the possible indexes a means summing over states really accessible to the particles; empty volumes V_a with $c_a = 0$ do not contribute to the sum. It is clear therefore that each f_a represents a possible state allowed for the diffusion system: the values f_a, $f_{a'}$, $f_{a''}$, ... in various points labeled by a, a', a'', ... quantify the ways of distributing the total mass M into various elementary volumes reached by the diffusing species during the diffusion process. Summing both sides of eq. 4,9 over the indexes a and b as done before, means therefore estimating the total probability of mass transport within the diffusion system; then let us introduce, even without carrying out any explicit calculation,

$$\sum_{a,b}\frac{J_a}{J_o} = -\sum_{a,b} f_a \log(f_a) + \\ + \sum_{a,b}\left(f_a \log(f_a+\gamma) - d_{ab}^2 \frac{\partial^2 \log(f)}{\partial x^2}\bigg|_{f_a}\right). \qquad (4,16)$$

Summing over all probabilities of diffusion paths, one finds the resulting configuration change of the diffusion system at any time. A few remarks are enough to guess what to expect from this equation. At $t \to 0$ one finds a sum of terms $f_a \log(1+\gamma/f_a)$, which for $\gamma \to 0$ tend to γ, plus terms that contain the factor d_{ab}; since in this limit $x_a - x_b \to 0$, neither of them contributes to the sum. At $t > 0$ both addends contribute to the sum. At the equilibrium asymptotic time where again $\gamma = 0$ the sum vanishes according to eq. 4,15 because $f_a \to f^{eq}$ everywhere; this result agrees with the statistical limit $\sum_{a,b} J_a/J_o = 0$ previously inferred, which actually is the macroscopic result revealed by the experience. The first addend at right hand side is clearly an entropic term, whereas f_a defined in eq. 4,5 must have the probabilistic significance of thermodynamic state related to the current configuration of the diffusion system. In effect it is possible to define the limit value f^{eq} such that $\sum_{a,b}(f^{eq}) = 1$ whatever the number of terms of the sum might be; indeed according to eq. 4,12 the finite limit γ_{ab} for $c_b \to c_a$ and $\gamma \to 0$ has been defined finite but not specified; the value of γ_{ab} can be therefore taken as that fulfilling the required property of f^{eq}. If so the first sum of eq. 4,16 is such that when the system evolves towards the equilibrium then

$$-\sum_{a,b} f_a \log(f_a) \to -\sum_{a,b} \log(f^{eq}).$$

The possibility of relating f_a to the thermodynamic prob-

ability of states allowed to the diffusing particles defines the physical meaning of the time parameter γ: depending on the value of this latter the totality of possible values of x_a and x_b, whatever they might be, corresponds to a possible arrangement of diffusing particles at the current time starting from an arbitrary initial configuration in the diffusing medium. According to eq. 4,10 it appears that $\gamma = 0$ at $t = 0$ defines the initial configuration. So, through the totality of possible paths from any x_b to any x_a, the parameter $\gamma > 0$ provides an indication of the order→disorder evolution of the configuration of the diffusion system as a function of time. Rewrite now eq. 4,16 as follows

$$\sum_{a,b} \frac{J_a}{J_o} = \frac{S_t}{k_B} - \frac{S_o}{k_B}, \tag{4,17}$$

where

$$\frac{S_t}{k_B} = -\sum_{a,b} f_a \log(f_a), \tag{4,18}$$

$$\frac{S_o}{k_B} = -\sum_{a,b} \left(f_a \log(f_a + \gamma) - d_{ab}^2 \frac{\partial^2 \log(f)}{\partial x^2}\bigg|_{f_a} \right).$$

The ratio J_a/J_o has been previously identified as the local chance of net mass flow between two arbitrary points of the diffusion system; the sum at left hand side is therefore the flow efficiency throughout the whole diffusion system, i.e. $\Pi_{netflow} = \sum_{a,b} J_a/J_o$. It is possible therefore to introduce the total chance of mass transfer, Π_{tr}, with and without net mass flow such that of course $\Pi_{tr} = \Pi_{netflow} + \Pi_{nonetflow}$ with obvious notation. This kind of definition is suggested by the possibility of normalizing Π_{tr} to 1. Hence comparing with eqs. 4,17 and 4,18 one infers

$$\Pi_{tr} = \frac{S_t}{k_B}, \qquad \Pi_{nonetflow} = \frac{S_o}{k_B}.$$

Of course S_t, the most general statistical definition of entropy, is also the most general way to describe the configuration of N diffusing particles in the N_V volumes available in the diffusion system, regardless of whether or not the configuration entails a net displacement of matter; instead S_o, which does not refer to net transfer of atoms, counts simply the number of ways to arrange any prefixed distribution of particles and thus the thermodynamic probability of any configuration. Hence the entropic terms concern two different kinds of diffusion mechanisms allowed to occur as a function of time. In effect the possibility that $x_b \to x_a$ is not excluded in the present model even at times t_1, t_2,..; it would be enough to define γ for instance as $t(t-t_1)(t-t_2)/(t+t_o)^4$ in agreement with the previous considerations at $t \to 0$ and $t \to \infty$ and at any time where $x_b \to x_a$ entails $c_b \to c_a$ too. Further considerations are possible about the results hitherto obtained.

5 Discussion

The eqs. 3,1 only have been exploited to highlight the link between concentration gradient law and entropy of diffusion system through elementary considerations. Both concepts have been extracted through elementary algebraic manipulations of the left and right hand sides of the unique eq. 4,3. No hypotheses "ad hoc" have been introduced about the physical features of the diffusion system and its driving mechanisms, leading for instance to Markovian jumps or not, interstitial or defect activated jumps, collisions in gas phase and so on. This is due to the general worth of eqs. 3,1 regardless of the specific system concerned: the present conclusions hold in principle for diffusion processes in solid or liquid or gas phase.

Regarding the statistical formulation of the uncertainty as fundamental principle of nature, the diffusion particles result randomly delocalized within elementary volumes $V = \Delta x^3$ into which can be ideally subdivided the whole system, whose size is however inessential to infer the entropic terms $-f \log f$; these volumes control the concentrations c, which in turn define the thermodynamic states allowed to the diffusing particles in relation to their occupation probability. No assumption was made about the coordinates of the points x_a and x_b falling within the respective elementary volumes, whose number, size and position indeed have been never specified in section 4. In fact such a kind of local information is irrelevant to calculate the entropy; it is enough to compute how N particles can be distributed in N_V volume elements, regardless of how many and where these latter might actually be in the diffusion medium. For this reason the model describes the time evolution of the whole system even without knowing in detail how is progressively modified the configuration of particles and volumes as a function of time. Actually eqs. 4,17 admit also empty elementary volumes that however do not contribute to the total entropy of the system, in fact determined by the distribution of particles only. So S_o in eq. 4,17 corresponds in general to the ways of distributing particles into available microstates described by Δx^3, possibly taking into account the indistinguishability of identical particles, through a dynamical pattern of particles exchanging their occupation volumes even without net mass flow. In effect, also this kind of information does not require a detailed knowledge on the local motion of particles. Nothing is known about this motion within their own Δx^3, because it would require some sort of local information about x and p_x. Being impossible to establish if within this arbitrary volume the motion is for instance Markovian or not, one must admit that both chances are in fact allowed; this also justifies why the diffusing species is involved in mass transfer process with and without net displacement of particles.

This conclusion does not conflict with the fact that J_x entails explicitly an effective concentration gradient; eq. 4,4 is simply the differential formulation of a physical law related

to the driving force that triggers the displacement, see eqs. 4,1 and 2,3 and 2,5 as well. The quantum approach behind this step accounts for the physical basis of eq. 2,2, whereas the definition 2,1 has now the rank of a corollary of eq. 4,4 rather than a mere definition: now the physical dimensions of eq. 2,1 are required by quantum motivations, rather than being suggested by a reasonable assumption. Indeed the available information about the diffusion system is inferred in the typical way of quantum mechanics, i.e. without requiring an exact local knowledge about position and momentum of the particles, as follows:

(i) from a macroscopic point of view, through J_o of eq. 4,8 and the entropic terms of eq. 4,17;

(ii) through the probabilistic meaning of the ratio J_a/J_o, which indeed represents the probability of effective mass transport as concerns the chances of Marcovian or non-Marcovian displacements.

Non-trivial consequence of these constrains about our degree of information is the heuristic achievement resulting from the quantum approach with respect to that provided by the classical physics where, from the point of view of the continuity equation, the general character of both Fick's laws is merely due to the lack of sinks/sources perturbing the diffusion process. This fact appeared already in the classical section 2 when it was found that in general $\mathbf{F} \neq \mathbf{F}'$, i.e. the driving force controlling the mass transport is in principle different from that due to local perturbations; the former was uniquely inferred from general hypotheses, eqs. 2,1 and 2,2, the latter remained instead unspecified and does so still now. This is not incompleteness of the present model, but rather the statement that the local perturbations must be purposely specified case by case depending on the physical features of the diffusion system. The worth of any theoretical approach depends on its ability to be generalized beyond the specific problem for which it was formerly conceived. In the case of diffusion the generalization is evident: several important physical laws are expressed through the gradient of a well defined function.

One example is the Fourier equation, $\mathbf{J}_Q = K\nabla T$, where K is the heat conductivity and $\mathbf{J}_Q$ the heat flow; also the Ohm law, $\mathbf{I} = R^{-1}\nabla V$, exhibits a similar form involving the electrical resistance R and the electric potential V to describe the displacement of charges per unit time. Although a common gradient law describes in the former case the transport of heat and in the latter that of electrons, both equations involve forms of kinetic energy, respectively due to the oscillation frequency of atoms/ions/molecules within the heat diffusion thermodynamic system and to the velocity of electrons propagating within a conductor. The entropic aspects in these systems are clear. In the former case they were already evidenced by the crucial Boltzmann intuition, although in lack of any quantum reference; it is not surprising that indeed the statistical definition of entropy inferred here goes back to the early times when the thermodynamics was essentially the science of heat exchanges. The entropy difference in the absence and presence of an electric field is also evident in the latter case: without electric field the motion of the electrons is random, in the $\mathbf{k}$ space it is represented by a sphere; the presence of the field instead orients the motion of the electrons along a preferential direction. The applied field triggers thus a more ordered motion of electrons, which suggests in turn a loss of total entropy. The analogy with the case discussed in section 4 is clear, although the respective entropy changes have opposite sign. This is not surprising: in an isolated system the entropy always increases, in a system interacting with an external field this is not necessarily true. In all cases however the gradient-like laws, mass diffusion, heat diffusion and Ohm law, are similarly consistent with entropic terms describing the actual numbers of accessible states during the displacement of matter or energy. Another consequence of the generality of the present model concerns the driving force of the diffusion process. In section 2, eq. 2,5 was inferred from eqs. 2,1 and 2,2, the only equations available. Of course the same can be done identically here, though on a more profound quantum basis. Yet the approach carried out in section 4 allowed inferring eq. 4,1, which introduces the concept of force directly as a consequence of eqs. 3,1 and deserves thus further considerations.

First of all, the quantum nature of the mass flow can be evidenced replacing v_x of eq. 4,1 into the x-component of eq 2,1, which yields thanks to eq. 2,5

$$J_x = \frac{k_B T}{n\hbar}\Delta x^2 c \frac{\partial \log(c/c_o)}{\partial x}. \qquad (5,1)$$

So, simply identifying F_x of eq. 2,5 with that of eq. 4,1 appear again terms of J_x having the form $c\partial\log(c)/\partial x$, which can be handled in a completely analogous way as in section 4 to infer entropic terms like $c\log(c/c_o)$ of eq. 4,5. Moreover $J_x \to 0$ for $n \to \infty$ agrees with eq. 4,16; an increase of entropy due to the increase of states accessible to the diffusion system corresponds to the reaching of asymptotic equilibrium where the net mass flow vanishes. As expected, the result obtained via the time coordinate defining v_x agrees with that previously obtained through the space coordinate only. Yet it is worth remarking that the combined information of the first eq. 4,18 plus eq. 5,1 regards this time behavior of any isolated diffusion system as a spontaneous evolution process: indeed $t \to \infty$ requires $J_x \to 0$ that in turn requires a maximum number of allowed states $n \to \infty$. Two fundamental statements of thermodynamics appear here as corollaries of eqs. 3,1: the statistical formula of entropy and the entropy increase in an isolated system.

Let us exploit eq. 5,1 noting that $k_B T/n\hbar$ has physical dimensions of time. So compute this equation at the time τ where the total diffusion spread lies within an average value of Δx^2 computed starting from $\Delta x^2 \to 0$ at $t = 0$ up to the value $\Delta x^2 = \Delta x_\tau^2$ at the time τ; this means assigning to Δx^2 the particular mean value $\overline{\Delta x^2} = \Delta x_\tau^2/2$ averaged between

zero and Δx_τ^2. Comparing with eq. 4,4, one finds immediately the known Einstein's one-dimensional result

$$D = \frac{1}{2}\frac{\Delta x_\tau^2}{\tau}.$$

6 Heuristic aspects of the quantum uncertainty

The present section, based on wide-ranging considerations about v_x, extends somewhat the preliminary remarks introduced in section 3 and has prospective worth. The aim is to emphasize that F_x of eq. 4,1 has actually a physical meaning much more general and contains much more information than the mere eq. 2,5. The byproduct of eqs. 3,1 proposed here is so short, straightforward and relevant to deserve being sketched although, strictly speaking, beyond the mere purposes of the present model; accordingly, however, the results hitherto inferred appear as a particular kind of selected physical information extracted from a broader context able to link topics apparently dissimilar.

Key tools of the following considerations are the replacements 3,2 that compel changing the way to formulate any physical property P from the usual form $P(x, p_x, t)$ to $P(\Delta x, \Delta p_x, \Delta t)$ and thus to $P(\Delta x, n, \Delta\varepsilon)$. In effect the paper [4] has shown that the number n of states coincides with the quantum number appearing in the eigenvalues of the harmonic oscillator, while the papers [5, 6] show that this is true in general; e.g. the number l of states calculated for the angular momentum coincides with the orbital quantum number. The first remark concerns the two ways of expressing F_x in eqs. 4,1:

(i) F_x follows from the definition of momentum itself, $\Delta\dot{p}_x = m\dot{v}_x$, and involves directly the mass m, previously introduced with mere reference to the concentration of diffusion particles and now regarded in general as the mass of any particle accelerated in Δx;

(ii) $F_x = -n\hbar\Delta x^{-2}v_x$ does not involve directly any mass but the deformation rate, $\Delta\dot{x} = v_x$, of Δx only.

Why in (ii) the mere time deformation of Δx in the phase space surrogates the presence of an accelerated mass? The answer rests on the same considerations already introduced in section 4: if a growing/shrinking range is accessible to a particle, then this latter can move faster/slower while being still therein delocalized; the fact that the particle can accelerate/decelerate simply reaffirms once more that nothing in known about how any dynamical variables change within the respective delocalization ranges.

However, in lack of constraining hypotheses, there is no reason to exclude that this idea holds regardless of whether the range sizes are stationary or not. Otherwise stated: slow motion in a short range or faster motion in a larger range are two indistinguishable chances, both allowed to occur for a particle by the lack of local information inherent the eqs. 3,1 and in fact both occurring. This rationalizes why just the uncertainty of x, p_x, ε and t links the deformation rate of time dependent range sizes of the phase space to the acceleration of any particle, possible and thus actual. The size and position of any range require a reference system to be defined in principle, although never quantifiable.

Consider for instance $\Delta x = x_t - x_o$ and $\Delta p_x = p_t - p_o$: the coordinate x_o, whatever it might be, is defined with respect to the origin O of an arbitrary reference system R, while the same also holds for the momentum p_o of the range Δp_x conjugate to Δx. So a free particle is described in R by its own Δx and Δp_x; indeed eqs. 4,1 have been inferred in R keeping constant x_o and regarding x_t as a time function. Yet, if neither of these boundaries is specifiable, one could also think x_t fixed and x_o time function. The difference is apparent: the displacement of x_o means that now Δx deforms while contextually moving in R, as O displaces at rate $-\partial x_o/\partial t$ with respect to x_t. Thus it is possible to introduce another reference system R_o solidal with x_o such that a particle accelerated in R is at rest in R_o, which moves with the same acceleration in R. Clearly still acts on the particle a force that justifies the acceleration of R_o in R, although however the particle is in fact at rest in R_o.

The conclusion of this reasoning is well known: a particle at rest in an accelerated reference frame is subjected to a force F_x indistinguishable from that due to the presence of mass. Of course with large sized Δx one can speak about average force F_x, whereas in a small sized range F_x takes a value better and better defined. This statement is nothing else but the equivalence principle, here inferred as a corollary of eqs. 3,1. After having introduced in eqs. 4,1 $F_x = m\dot{v}_x$, can be inferred also the link between F_x and Newton's law after these preliminary remarks? Of course let us start again from eqs. 4,1 with v_x and $\dot{v}_x$ defined in any R.

First of all, the fact that the mass in eq. 4,1 is unique and that the equivalence principle has been obtained elaborating independently both sides of $m\dot{v}_x = F_x = -n\hbar\Delta x^{-2}v_x$ shows the identity of inertial and gravitational mass. Moreover just the fact the unique mass m must somehow appear also in the second equality compels putting $v_x = \zeta' m$ via an appropriate dimensional factor ζ'; hence $F_x = -n\hbar\zeta' m\Delta x^{-2}$ with the acceleration no longer appearing explicitly in this expression, which rather has the form of an interaction force $F_x^{m,\zeta'}$ between m and another entity that can be nothing else but ζ'.

This result suggests a more interesting form of F_x putting $\zeta' = n\sum_k \zeta_k m'^k$, being ζ_k coefficients of the power series development of ζ' and m' a further arbitrary mass that interacts with m. The series truncated at the first order only yields approximately $\zeta' \simeq n\zeta m'$, with ζ unique proportionality factor; here n is inessential and does not play any role because, being m' arbitrary, $m'n$ is another value arbitrary as well. In this way one finds $F_x^{m,m'} \approx -\hbar\zeta(m/\Delta x)(m'/\Delta x)$ at the first order of approximation, i.e. an attractive force is originated between the linear densities $m/\Delta x$ and $m'/\Delta x$ of masses by definition delocalized within Δx. This sensible result appears better under-

standable thinking to particle waves that propagate through Δx rather than to point particles moving randomly within Δx.

Moreover the proportionality factor ζ can be regarded as a constant since the arbitrary masses m and m' account for the arbitrariness of v_x. With the notation $\zeta = G/\hbar$ one recognizes the approximate Newton law; the classical distance $x_{m,m'}$ between local coordinates exactly known of particles is replaced by any random distance falling within the uncertainty range including them.

Obviously Δx^{-2} shows that the functional dependence of $F_x^{m,m'}$ on all possible distances between the masses is like $x_{m'm}^{-2}$. This confirms that effectively the diffusion particles are acted by the force F_x, whose physical meaning can be extended even to the gravitational interaction. Note however that actually both signs are allowed for the velocity component v_x along x, which correspond to the signs of $\partial\Delta x/\partial t$ depending on whether Δx shrinks or expands as a function of time.

In agreement with the idea of phase space-time deformation in the presence of mass, one would expect thus $v_x = \pm\zeta' m$, i.e. even a negative value of m. This conclusion emphasizes nothing else but the existence of antimatter. After this instance about how eqs. 3,1 can be purposely exploited, let us proceed with another example short enough to be mentioned here, i.e. the Coulomb law. It is not a chance that even this latter has a form similar to that of the Newton law, with the charges playing the role of the masses. To emphasize the reason of this similarity, let us introduce in eq. 4,1 the fine structure constant $\alpha = e^2/\hbar c$. Eliminating $\hbar$ eq. 4,1 reads $F_x = e'e/\Delta x^2 = m_e a_x$, where now m_e is the electron mass and $e' = nv_x(c\alpha)^{-1}e$. This latter reads more expressively $e' = \pm n\,|v_x|\,(c\alpha)^{-1}e$. Again, the charges interact through their linear densities $e/\Delta x$ and $e'/\Delta x$ for the reasons previously explained. Also the electron charges appear therefore because of the phase space-time deformation in the presence of the mass m_e. Once more is crucial the characteristic value of v_x of charged particles; for instance $v_x = 0$ would describe a neutral particle, whereas it also appears that a massless particle would be chargeless as well. A boundary condition of the problem is that for an appropriate value n^* of the integer n one must find $e' = e$, as nothing hinders indeed just such a possibility. So $e' = \pm(n/n^*)e$; e.g. for a couple of electrons one must take $n = n^*$ i.e. $|v_x| = c\alpha$, whose value seems therefore to be a combined constant of nature. It is reasonable to assume $n^* = 3$ since actually one should consider v_x, v_y and v_z for the respective components replacing the early F_x, for simplicity the only one hitherto considered, whereas the number n of states should be counted as $n = n_x + n_y + n_z$. Take the ground values $n_x = n_y = n_z = 1$ and consider the three chances $v_x \neq 0, v_y \neq 0, v_z \neq 0$ and $v_x \neq 0, v_y \neq 0$ and $v_x \neq 0$ only. This means considering the charges of particles resulting from $n = 1, 2, 3$ with $n^* = 3$. As inferred before, $n = n^*$ holds for protons and electrons. Yet, in addition to $e' = \pm e$, possible values of e' result respectively to be $e' = \pm e/3$ or $e' = \pm(2/3)e$ as well, i.e. particles with fractional charges should also exist in nature. But, being n arbitrary, what about hypothetical charges described by $n > n^*$?

A full discussion on this question is clearly far beyond the purpose of the present paper; further work is in progress on this specific topic. As concerns the results hitherto introduced, it is enough to conclude that the formal analogy between the Fourier law and the mass/charge transport laws is due to their common quantum basis, discussed here with reference to the entropic aspects too, that goes back to the intimate quantum nature of the entropy and Newton and Coulomb forces themselves.

7 Conclusion

The quantum origin of the diffusion law has been described with the help of eqs. 3,1 only. The assumption of uncertainty that allowed to calculate the energy levels of many-electron atoms and diatomic molecules, enables even the basic law controlling the transport of matter to be inferred in a very simple way. It is also remarkable that elementary considerations on eqs. 3,1 open the way to further results much more general than the specific task to which they were initially addressed in the present diffusion model. This emphasizes the heuristic character of eqs. 3,1: the uncertainty, regarded itself as a fundamental principle of nature rather than as mere corollary of commutation rules of operators, appears a key tool to infer a conceptual background unifying seemingly different physical phenomena. As concerns the present model, the level of comprehension provided by the approach based on the phase space-time uncertainty opens the way to more specific considerations on the possible mechanisms of transport in solid, liquid and gas phases.

Submitted on September 1, 2011 / Accepted on September 19, 2011

References

1. Heitjans P., Karger J. (Eds.) Diffusion in condensed matter: Methods, Materials, Models, 2nd edition, Birkhauser, (2005).
2. Incropera F., DeWitt P.D.P., Bergman T.L., Lavine A.S. Fundamentals of Heat and Mass Transfer, 6th edition, John Wiley, (2006).
3. Mehrer H. Diffusion in solids, Springer, Berlin, (2007).
4. Tosto S. An analysis of states in the phase space: the anharmonic oscillator. *Progress in Physics*, 2011, v. 4, 29–36.
5. Tosto S. An analysis of states in the phase space: the energy levels of quantum systems, *Nuovo Cimento B*, 1996, v. 111 (2), 193–215.
6. Tosto S. An analysis of states in the phase space: the diatomic molecules. *Nuovo Cimento D*, 1996, v. 18 (12), 1363–1394.

Dynamical Space: Supermassive Galactic Black Holes and Cosmic Filaments

Reginald T. Cahill and Daniel J. Kerrigan

School of Chemical and Physical Sciences, Flinders University, Adelaide 5001, Australia.
E-mail: Reg.Cahill@flinders.edu.au, Daniel.Kerrigan@flinders.edu.au

The unfolding revolution in observational astrophysics and cosmology has lead to numerous puzzles: "supermassive" galactic central black holes, galactic "dark matter" halos, relationships between these black hole "effective" masses and star dispersion speeds in galactic bulges, flat spiral galaxy rotation curves, cosmic filaments, and the need for "dark matter" and "dark energy" in fitting the Friedmann universe expansion equation to the supernovae and CMB data. Herein is reported the discovery of a dynamical theory for space which explains all these puzzles in terms of 3 constants; G, α - which experimental data reveals to be the fine structure constant $\alpha \approx 1/137$, and δ which is a small scale distance, perhaps a Planck length. It is suggested that the dynamics for space arises as a derivative expansion of a deeper quantum foam phenomenon. This discovery amounts to the emergence of a unification of space, gravity and the quantum.

1 Dynamical Space

The many mysteries of cosmology, such as supermassive galactic black holes, cosmic filaments, "dark matter" galactic haloes, flat spiral-galaxy rotation curves, "dark energy" effects in expansion of the universe, and various unexplained correlations between galactic black hole masses and star velocities, all suggest that we have an incomplete account of space and gravity. We report herein the discovery of such a theory and its successful testing against the above phenomena, and as well against laboratory and geophysical gravity experiments. If space is, at a deep level, a quantum system, with dynamics and structure, then we expect a derivative expansion would give a classical/long-wavelength account. In the absence of that quantum theory we construct, phenomenologically, such an account in terms of a velocity field [1]. In the case of zero vorticity we obtain

$$\nabla\cdot\left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v}\right) + \frac{\alpha}{8}\left((trD)^2 - tr(D^2)\right) + \\ + \frac{\delta^2}{8}\nabla^2\left((trD)^2 - tr(D^2)\right) + ... = -4\pi G\rho$$

$$\nabla\times\mathbf{v} = \mathbf{0}, \quad D_{ij} = \frac{\partial v_i}{\partial x_j} \tag{1}$$

where the major development reported herein is the discovery of the significance of the new $\delta-$term, with δ having the dimensions of a length, and presumably is the length scale of quantum foam processes. This term is shown to be critical in explaining the galactic black hole and cosmic filament phenomena. This δ is probably a Planck-like length, and points to the existence of fundamental quantum processes. If $\delta = 0$ (1) cannot explain these phenomena: δ must be non-zero, no matter how small, and its value cannot be determined from any data, so far. G is Newton's constant, which now appears to describe the dissipative flow of quantum foam into matter, and α is a dimensionless self-coupling constant, that experiment reveals to be the fine structure constant, demonstrating again that space is fundamentally a quantum process. We briefly outline the derivation of (1). Relative to the non-physical classical embedding space, with coordinates $\mathbf{r}$, and which an observer also uses to define the velocity field, the Euler constituent acceleration of the quantum foam is

$$\mathbf{a} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} \tag{2}$$

and so, when $\alpha = 0$ and $\delta = 0$, (1) relates this acceleration to the density of matter ρ, and which will lead to Newton's account of gravity. The matter acceleration is found by determining the trajectory of a quantum matter wavepacket. This is most easily done using Fermat's maximum proper-travel time τ:

$$\tau = \int dt\sqrt{1 - \frac{\mathbf{v}_R^2(\mathbf{r}_0(t), t)}{c^2}} \tag{3}$$

where $\mathbf{v}_R(\mathbf{r}_o(t), t) = \mathbf{v}_o(t) - \mathbf{v}(\mathbf{r}_o(t), t)$, is the velocity of the wave packet, at position $\mathbf{r}_0(t)$, wrt the local space. This ensures that quantum waves propagating along neighbouring paths are in phase, and so interfere constructively. This maximisation gives the quantum matter geodesic equation for $\mathbf{r}_0(t)$

$$\mathbf{g} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} + (\nabla\times\mathbf{v})\times\mathbf{v}_R \\ - \frac{\mathbf{v}_R}{1 - \dfrac{\mathbf{v}_R^2}{c^2}}\frac{1}{2}\frac{d}{dt}\left(\frac{\mathbf{v}_R^2}{c^2}\right) + ... \tag{4}$$

with $\mathbf{g} \equiv d\mathbf{v}_o/dt$. The 1st term in $\mathbf{g}$ is the Euler space acceleration $\mathbf{a}$, the 2nd term explains the Lense-Thirring effect, when the vorticity is non-zero, and the last term explains the precession of orbits. In the limit of zero vorticity and neglecting

relativistic effects (1) and (4) give

$$\nabla \cdot \mathbf{g} = -4\pi G\rho - 4\pi G\rho_{DM}, \quad \nabla \times \mathbf{g} = \mathbf{0} \tag{5}$$

where

$$\rho_{DM} \equiv \frac{\alpha}{32\pi G}\left((trD)^2 - tr(D^2)\right) + \frac{\delta^2}{32\pi G}\nabla^2\left((trD)^2 - tr(D^2)\right) + ... \tag{6}$$

This is Newtonian gravity, but with the extra dynamical terms which has been used to define an effective "dark matter" density. This ρ_{DM} is not a real matter density, of any form, but is the matter density needed within Newtonian gravity to explain dynamical effects caused by the α and δ-terms in (1). It is purely a space/quantum-foam self-interaction effect. Eqn.(3) for the elapsed proper time maybe written in differential form as

$$d\tau^2 = dt^2 - \frac{1}{c^2}(d\mathbf{r}(t) - \mathbf{v}(\mathbf{r}(t), t)dt)^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \tag{7}$$

which introduces a curved spacetime metric $g_{\mu\nu}$ for which the geodesics are the quantum matter trajectories when freely propagating through the quantum foam. When $\alpha = 0$ and $\delta = 0$, and when ρ describes a sphere of matter of mass M, (1) has, external to the sphere, a static solution $\mathbf{v}(\mathbf{r}) = -\sqrt{2GM/r}\hat{\mathbf{r}}$, which results in Newton's matter gravitational acceleration $\mathbf{g}(\mathbf{r}) = -GM/r^2\hat{\mathbf{r}}$. Substituting this $\mathbf{v}(\mathbf{r})$ expression in (7), and making the change of time coordinate

$$t \to t' = t - \frac{2}{c}\sqrt{\frac{2GMr}{c^2}} + \frac{4GM}{c^3}\tanh^{-1}\sqrt{\frac{2GM}{c^2 r}}, \tag{8}$$

(7) becomes the standard Schwarzschild metric, and which is the usual explanation for the galactic black hole phenomenon, see [3–5], namely a very small radius but very massive concentration of matter. To the contrary we show here that the observed galactic black holes are solutions of (1), even when there is no matter present, $\rho = 0$. These solutions are quantum foam solitons.

The above $\mathbf{v}(\mathbf{r}) = -\sqrt{2GM/r}\hat{\mathbf{r}}$ solution also explains why the $\alpha-$ and $\delta-$terms in (1) have gone unnoticed, namely that for these solutions $(trD)^2 - tr(D^2) = 0$. It is for this reason that the $\alpha-$ and $\delta-$terms are now included, namely that Newton's inverse square law for gravity is preserved for solar system situations, and from which Newton determined his theory from Kepler's analysis of Brahe's planetary data. The key point is that the solar system is too special to have revealed the full complexity of the phenomenon of gravity.

However just inside a planet the $\alpha-$term becomes detectable, and it results in the earth's matter acceleration g being slightly larger than that predicted by Newtonian gravity, and we obtain from (1)

$$\Delta g = g_{NG}(d) - g(d) = -2\pi\alpha G\rho(R)d + O(\alpha^2), \quad d > 0 \tag{9}$$

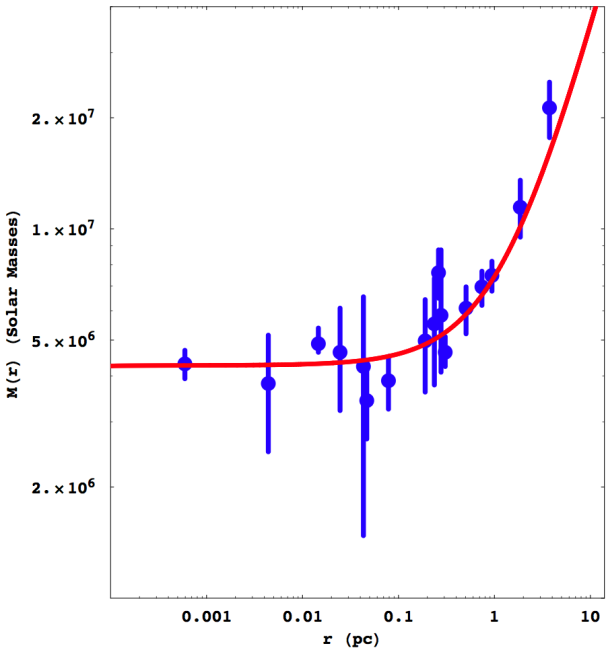


Fig. 1: The $M(r)$ data for the Milky Way SgrA* black hole, showing the flat regime, that mimics a point-like mass, and the rising form beyond $r_s = 1.33$pc, as predicted by (12), but where M_0 and r_s parametrise a quantum foam soliton, and involves no actual matter. The left-most data point is from the orbit of star S2 - using the Ghez *et al.* [3] value $M_0 = 4.5 \pm 0.4 \times 10^6$ solar masses. The other data is from Camenzind [5], but which requires these remaining data points to be scaled up by a factor of 2, presumably arising from a scaling down used to bring this data into agreement with a smaller initial value for M_0.

down a bore hole at depth d. This involves only α as the δ-term is insignificant near the surface. The Greenland Ice Shelf bore hole data [6] and Nevada bore hole data [7], both give $\alpha \approx 1/137$ to within observational errors, even though the ice and rock densities $\rho(R)$ differ by more than a factor of 2 [2]. So this result for α is robust, and shows that α is the fine structure constant $\alpha = e^2\hbar/c$, with α probably the more fundamental constant, and now showing up in the quantum foam account for gravity. As well laboratory measurements of G, modified Cavendish experiments, have always shown anomalous and inconsistent results [10, 11], revealing a systematic effect not in Newtonian gravity. Indeed the Long 1976 laboratory experiment to measure G, reported the anomaly to have magnitude $\delta_L = 0.0037 \pm 0.0007$ [8] (this δ_L is not related to δ in (1)), which equals $0.5/(136 \pm 26)$, showing that α can be measured in laboratory gravity experiments, of the type pioneered by Long.

1.1 Black Holes and Filaments as Quantum Foam Solitons

For the special case of a spherically symmetric flow, and in the absence of matter $\rho = 0$, we set $\mathbf{v}(\mathbf{r}, t) = \hat{\mathbf{r}}v(r)$. Then (1) has exact static two-parameter, v_0 and $\kappa \geq 1$, solutions

$$v(r)^2 = v_0^2(\kappa - 1)\frac{\delta}{r}\left(1 - {}_1F_1\left[-\frac{1}{2} + \frac{\alpha}{4}, -\frac{1}{2}, -\frac{r^2}{\delta^2}\right]\right) + \\ + v_0^2\kappa\left(\frac{4-2\alpha}{3}\right)\frac{r^2}{\delta^2}\frac{\Gamma(\frac{2-\alpha}{4})}{\Gamma(-\frac{\alpha}{4})}\,{}_1F_1\left[1 + \frac{\alpha}{4}, \frac{5}{2}, -\frac{r^2}{\delta^2}\right], \tag{10}$$

where ${}_1F_1\,[a, b, w]$ is the confluent hypergeometric function. Here v_0 is a speed that sets the overall scale, and κ is a structural parameter for the black hole, and sets the relative significance of the two terms in (11) and (12), and which is determined by the history of the black hole: in-falling matter increases κ, and values of both v_0 and κ are affected by surrounding matter if $\rho \neq 0$. In the limit $r \gg \delta$

$$v(r)^2 \approx A\frac{\delta}{r} + B\left(\frac{\delta}{r}\right)^{\alpha/2}. \tag{11}$$

However $v(r) \to 0$ as $r \to 0$ when $\delta \neq 0$, and so the δ-term dynamics self-regulates the interior structure of the black hole, which has a characteristic radius of $O(\delta)$. Inside this radius the in-flow speed goes to zero, and so there is no singularity. Hence there is a naturally occurring UV cutoff mechanism. Eqn. (**??**) gives an asymptotic form for $g(r)$, which is parametrised by an "effective mass" $M(r)$ within radius r: $g(r) = GM(r)/r^2$. In terms of observable $M(r)$ (11) gives a two-parameter description

$$M(r) = M_0 + M_0\left(\frac{r}{r_s}\right)^{1-\alpha/2} \tag{12}$$

r_s is the distance where $M(r_s) = 2M_0$. $M(r)$ from the Milky Way SgrA* black hole [3–5] is shown in Fig.1, and the best fit gives $r_s = 1.33$pc. This remarkable data comes from observations of orbits of stars close to SgrA*, and in particular the star S2, which has an elliptical orbit with a period of 15.2±0.11 years, and is the left-most data point in Fig.1. This dynamical space solution exhibits an effective point-like mass acceleration for $r < r_s$, where $M(r)$ is essentially constant, and for $r > r_s$ an increasing $M(r)$. At the outer-most data point the presence of stars within the galactic core begin to become apparent, with $M(r)$ becoming larger than the form predicted in (12). Note that if $\delta = 0$, then the flat feature in $M(r)$ is absent, while if $\alpha = 0$ the rise in $M(r)$ is absent, and the flat feature continues outwards. Intriguingly then the role of the $\delta-$term dynamics is critical to the effective point-like mass description of the inner region of the black hole, even though there is no actual matter present. It is this region of $M(r)$ that explains the inner star elliptical orbits - with $\delta = 0$ the $\alpha-$term produces a "weak" black hole, but with $g(r) \sim 1/r^{1+\alpha/2}$, which does not produce the observed star orbits. Eqn. (12) is in terms of observables. If we best-fit the data using an $M(r)$ directly from (10), by varying v_0, κ and δ, we find that there is no unique value of δ - v_0 and κ rescale to compensate for a deceasing δ, in the regime outside of the inner core to the black hole, but δ cannot be set to zero. This is evidence of the existence of a finite, but very small, structure to space, suggestive of a Planck-like fundamental length.

This black hole also explains the so-called "dark matter" halo. Asymptotically $\rho_{DM}(r)$ is related to the matter-less $M(r)$ via

$$M(r) = \int_0^r 4\pi r^2 \rho_{DM}(r) dr \tag{13}$$

giving

$$\rho_{DM}(r) = \frac{(1 - \alpha/2)M_0}{4\pi r_s^{1-\alpha/2} r^{2+\alpha/2}} \tag{14}$$

which decreases like $r^{-\gamma}$ with $\gamma = 2 + \alpha/2$. The value of the exponent γ has been determined by gravitational lensing for numerous elliptical galaxies in the Sloan Lens ACS Survey [12], and all give the generic result that $\gamma = 2$. Higher precision data may even permit the value of α to be determined. So the space dynamics completely determines ρ_{DM} in terms of observables M_0 and r_s.

Unlike the point-mass parametrisation of black holes, the above shows that the quantum foam black hole is an extended entity, dominating the galaxy from the inner regions, to beyond the central bulge, and even beyond the spiral arms. Indeed the $\rho_{DM}(r)$ in (14) predicts flat rotation curves, with orbital speed given by

$$v_{orb}^2(r) = GM_0\left(\frac{r_s}{r}\right)^{\alpha/2}\frac{1}{r_s} \tag{15}$$

but to which must be added the contribution form the matter density. For the Milky Way, we get the black hole contribution is $v_{orb} = 117$km/s at the location of the solar system, $r = 8$kpc, and determined by M_0 and r_s. That the black hole is an extended structure explains various observed correlations, such as that in [9] which reported a correlation between M_0 and the stellar speed dispersion in the bulge.

Eqn. (1), but only when $\delta \neq 0$, also has exact filament solutions

$$v(r)^2 = v_0^2\frac{r^2}{\delta^2}\,{}_1F_1\left[1 + \frac{\alpha}{8}, 2, -\frac{r^2}{2\delta^2}\right] \tag{16}$$

where r is the distance perpendicular to the axis of the filament, and $v(r)$ is the in-flow in that direction. In the limit $r \gg \delta$

$$v(r)^2 \sim 1/r^{\alpha/4} \;\text{ giving }\; g(r) \sim 1/r^{1+\alpha/4} \tag{17}$$

producing a long range gravitational attraction. Such cosmic filaments have been detected using weak gravitational lensing combined with statistical tomographic techniques. Again $v(r) \to 0$ as $r \to 0$ when $\delta \neq 0$, and so the δ-term dynamics self-regulates the interior structure of the filament, which has a characteristic radius of $O(\delta)$.

1.2 Expanding Universe

The dynamical 3-space theory (1) has a time dependent expanding universe solution, in the absence of matter, of the Hubble form $v(r,t) = H(t)r$ with $H(t) = 1/(1 + \alpha/2)t$, giving a scale factor $a(t) = (t/t_0)^{4/(4+\alpha)}$, predicting essentially a uniform expansion rate. This results in a parameter-free fit to the supernova redshift-magnitude data. In contrast the Friedmann model for the universe has a static solution - no expansion, unless there is matter/energy present. However to best fit the supernova data fictitious "dark matter" and "dark energy" must be introduced, resulting in the ΛCDM model. The amounts $\Omega_\Lambda = 0.73$ and $\Omega_{DM} + \Omega_M = 0.27$ are easily determined by best fitting the ΛCDM model to the above uniformly expanding result, without reference to the observational supernova data. But then the ΛCDM has a spurious exponential expansion which becomes more pronounced in the future.

2 Conclusions

The notion that space is a quantum foam system suggests a long-wavelength classical derivative-expansion description, and inspired by observed properties of space and gravity, such an effective field theory has been determined. This goes beyond the Newtonian modeling in terms of an acceleration field description - essentially the quantum foam is accelerating, but at a deeper level the acceleration is the Euler constitutive acceleration in terms of a velocity field. This velocity field has been detected experimentally, with the latest being from spacecraft earth-flyby Doppler shift data [13]. The dynamics of space now accounts for data from laboratory experiments through galactic black holes and filaments, to the expansion of the universe. We note that there is now no known phenomenon requiring "dark energy" or "dark matter". The black hole and cosmic filament phenomena require the existence of both α - the fine structure constant, and δ which is presumably a quantum foam characteristic Planck-like length scale. Gravity is now explainable as a emergent phenomenon of quantum foam dynamics, but only if we use as well a quantum wave description of matter - gravitational attraction is a quantum matter wave refraction effect, and also causes EM wave refraction. Hence the evidence is that we are seeing the unification of space, gravity and the quantum, pointing to a revolution in physics, and in our understanding of reality.

Submitted on September 1, 2001 / Accepted on September 19, 2011

References

1. Cahill R.T. in: Should the Laws of Gravitation be Reconsidered?, Múnera H.A. (Editor), Apeiron, Montreal, 2011, pp. 363–376.
2. Cahill R.T. Dynamical 3-Space: Emergent Gravity. *Progress in Physics*, 2011, v. 2, 44–51.
3. Ghez A.M., Salim S., Weinberg N.N., Lu J.R., Do T., Dunn J.K., Matthews K., Morris M., Yelda S., Becklin E.E., Kremenek T., Milosavljevic M., Naiman J. Measuring Distance and Properties of the Milky Way's Central Supermassive Black Hole with Stellar Orbits. *The Astrophysical Journal*, 2008, v. 689 (2), 1044–1062.
4. Gillessen S., Eisenhauer F., Trippe S., Alexander T., Genzel R., Martins F., Ott T. Monitoring the Stellar Orbits around the Massive Black Hole in the Galactic Center. *The Astrophysical Journal*, 2009, v. 692 (2), 1075–1109.
5. Camenzind M. Black Holes in Nearby Galaxies. GRK Vorlesung Würzburg, 2009.
6. Ander M.E., Zumberge M.A., Lautzenhiser T., Parker R.L., Aiken C.L.V., Gorman M.R., Nieto M.M., Cooper A.P.R., Ferguson J.F., Fisher E., McMechan G.A., Sasagawa G., Stevenson J.M., Backus G., Chave A.D., Greer J., Hammer, P., Hansen B.L., Hildebrand J.A., Kelty J.R., Sidles C., Wirt J. Test of Newton's inverse-square law in the Greenland ice cap. *Physical Review Letters*, 1989, v. 62, 985–988.
7. Thomas J., Vogel P. Testing the inverse-square law of gravity in boreholes at the Nevada Test Site. *Physical Review Letters*, 1990, v. 65, 1173–1176.
8. Long D.R. Experimental examination of the gravitational inverse square law. *Nature*, 1976, v. 260, 417–418.
9. Ferrarese L., Merrit D. A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies. *The Astrophysical Journal Letters*, 2000, v. 539, L9.
10. Reich, E.S. G-whizzes disagree over gravity. *Nature*, 2010, v. 466, 1030.
11. Davis R. Fundamental constants: Big G revisited. *Nature*, 2010, v. 468, 181–183.
12. Bolton A., Treu T., Koopmans L.V.E., Gavazzi R., Moustakas L.A., Burles S., Schlegel D.J., Wayth R. The Sloan Lens ACS Survey. V. The Full ACS Strong- Lens Sample. *Astrophysics Journal*, 2008, v. 682, 964–984.
13. Cahill R.T. Combining NASA/JPL One-Way Optical-Fiber Light-Speed Data with Spacecraft Earth-Flyby Doppler-Shift Data to Characterise 3-Space Flow. *Progress in Physics*, 2009, v. 4, 50–64.

A Generalized Displacement Problem in Elasticity

Peter Wilde
University of Applied Sciences Jena, Carl-Zeiss-Promenade 2, 07745 Jena, Germany.
E-mail: Peter.Wilde@TU-Ilmenau.de

By solving a special coupling boundary value problem for vector Helmholtz equations it is shown how the displacement boundary value problem in elasticity can be solved. It is shown that the generalized displacement problem possesses at most one solution.

1 Statement of the Problem

By D_i we denote a bounded domain in $\boldsymbol{R}^3$ with boundary S belonging to the class C^2, and by D_e the unbounded domain $D_e := \boldsymbol{R}^3 \setminus \overline{D_i}$. We assume that the normal vector n on S is directed into the exterior domain D_e. The physical meaning is that D_i is a fixed elastic solid with no volume forces present and D_e represents a homogeneous isotropic linear solid which is characterized by the density $\rho = 1$ (this is no loss of generality) and the Lamé parameters λ and μ. We consider time-harmonic elastic waves with circular frequency ω and it will be assumed that all Lamé constants and the frequency are positive. We assume that the elastic medium D_e is in welded contact to the rigid inclusion D_i, which means that we consider displacement boundary conditions.

To formulate the elasticity problems we introduce the following function spaces. By $C^{0,\alpha}(S)$ and $C_T^{0,\alpha}(S)$ we denote the spaces of Hölder continuous functions and Hölder continuous tangential fields ($0 < \alpha < 1$), respectively. The space $C_D^{0,\alpha}(S)$ denotes the subspace of Hölder continuous tangential fields possessing Hölder continuous surface divergence in the sense of the limit integral definition given by Müller [1]. Defining the differential operator $\Delta^* := \Delta + \frac{\lambda+\mu}{\mu}\operatorname{grad}\operatorname{div}$, where Δ is the Laplace operator and λ and μ are the Lamé elastic constants with $\mu > 0$ and $\lambda + 2\mu > 0$. For a positive frequency ω the wavenumbers κ_p and κ_s are defined by $\kappa_p := \omega/\sqrt{\lambda + 2\mu}$ and $\kappa_s := \omega/\sqrt{\mu}$. Now, the time-harmonic exterior displacement problem in elasticity can be formulated as

Problem D: Find a vector field $u \in C^2(D_e) \cap C(\overline{D_e})$ satisfying the time-harmonic elasticity equation

$$\Delta^* u + \kappa_s^2 u = 0, \text{ in } D_e, \tag{1}$$

the welded contact boundary conditions

$$u = f, \text{ on } S, \tag{2}$$

and the Sommerfeld radiation condition

$$(x, \operatorname{grad} u_j(x)) - i\kappa_j u_j(x) = o(\frac{1}{|x|}), \text{ for } |x| \to \infty,\ j = s, p, \tag{3}$$

uniformly for all directions $x := x/|x|$, where

$$u_p := \frac{-1}{\kappa_p^2}\operatorname{grad}\operatorname{div} u \text{ and } u_s := \frac{1}{\kappa_p^2}\operatorname{grad}\operatorname{div} u + u. \tag{4}$$

Here $f \in C^{0,\alpha}(S)$ is a given vector field.

By (a, b) and $[a, b]$ we denote the scalar product and vector product of the vectors a and b, respectively. The vector fields u_s and u_p are known as the rotational and irrotational parts of u, respectively. The rotational part corresponds to a dilatational or compressional wave and the irrotational part corresponds to a shearing wave. The wave numbers κ_s and κ_p are known as the slownesses of the rotational and irrotational waves, respectively.

That Problem D possesses at most one solution has already been discussed by Kupradze [2] and Ahner [3]. The existence of a solution has been shown by Hähner and Hsiao [4].

For any domain $D \subset \boldsymbol{R}^3$ with boundary ∂D we introduce the linear space of vector fields $u : D \to \boldsymbol{R}^3$ by

$$F(D) := \{u \mid u \in C^2(D) \cap C(\overline{D}),\ \operatorname{curl} u, \operatorname{div} u \in C(\overline{D})\ \}.$$

From the integral representation theorem for solutions of the time-harmonic elasticity equation, known as the Betti formulas [2], we see that the displacement field is analytic. Therefore, by using (4) u can be split into $u = u_p + u_s$. Differentiating both, u_p and u_s, we see that u_p is *curl*-free and that u_s is *divergence*-free. Furthermore, u_j is a solution of the vector Helmholtz equation $\Delta u_j + \kappa_j u_j = 0$, in D_e, for $j = s, p$.

This motivates us to study the following slightly more general coupling

Problem HD: Find two vector fields $u_s, u_p \in F(D_e)$ satisfying the vector Helmholtz equations

$$\left.\begin{array}{lll} \Delta u_s + \kappa_s^2 u_s = 0, & \text{in } D_e, & \kappa_s \neq 0,\ \Im(\kappa_s) \geq 0, \\ \Delta u_p + \kappa_p^2 u_p = 0, & \text{in } D_e, & \kappa_p \neq 0,\ \Im(\kappa_p) \geq 0, \end{array}\right\} \tag{5}$$

the coupling boundary conditions

$$\left.\begin{array}{lllll} [n, u_s] & + & [n, u_p] & = & c, \\ \operatorname{div} u_s & & & = & \gamma, \\ & & [[\operatorname{curl} u_p, n], n] & = & d, \\ (n, u_s) & + & (n, u_p) & = & \delta, \text{ on } S, \end{array}\right\} \tag{6}$$

and the radiation conditions

$$[\operatorname{curl} u_j, \hat{x}] + \hat{x} \operatorname{div} u_j - i\kappa_j u_j = o(1/|x|), \text{ for } |x| \to \infty, \tag{7}$$

and $j = s, p$, uniformly for all directions $\hat{x} := x/|x|$. Here $c \in C_D^{0,\alpha}(S)$ and $d \in C_T^{0,\alpha}(S)$ are given tangential fields and $\gamma, \delta \in C^{0,\alpha}(S)$ are given functions.

2 Uniqueness

By Problem HDS we denote the special case of Problem HD, with

$$\kappa_p^2 = \frac{\omega^2}{\lambda + 2\mu} \text{ and } \kappa_s^2 = \frac{\omega^2}{\mu}, \tag{8}$$

and the right-hand sides

$$\gamma = 0 \text{ and } d = 0.$$

Now we have the following equivalence

Theorem 3.1: 1) Let u be a solution of Problem D corresponding to the boundary data $f := n\delta - [n, c]$. Then

$$u_p := \frac{-1}{\kappa_p^2}\text{grad div}\, u \text{ and } u_s := \frac{1}{\kappa_p^2}\text{grad div}\, u + u,$$

is a solution of Problem HDS.

2) Let u_p, u_s be a solution of Problem HDS corresponding to the boundary data $c := [n, f]$, $\gamma = 0$, $d = 0$ and $\delta := (n, f)$. Then $u := u_p + u_s$ is a solution of Problem D.

Proof: We will show only part 2). Let u_p, u_s be a solution of Problem HDS corresponding to the boundary data $c := [n, f]$, $\gamma = 0$, $d = 0$ and $\delta := (n, f)$. Representing u_s via the representation theorem for solutions of the vector Helmholtz equation [6] it can be seen that $\text{div}\, u_s$ is a solution of the scalar Helmholtz equation $\Delta\text{div}\, u_s + \kappa_s^2\text{div}\, u_s = 0$ in D_e satisfying the homogeneous Dirichlet boundary condition $\text{div}\, u = 0$ and the Sommerfeld radiation condition. From the uniqueness theorem for the exterior Dirichlet problem [5, 6] we obtain $\text{div}\, u_s = 0$ in D_e.

Using the integral representation theorem for solutions of the vector Helmholtz equation [6] it can be seen that $\text{curl}\, u_p$ solves the vector Helmholtz equation $\Delta\text{curl}\, u_p + \kappa_p^2\text{curl}\, u_p = 0$ in D_e, fulfills the homogeneous electric boundary condition $[[\text{curl}\, u_p, n], n] = 0$ and $\text{div curl}\, u_p = 0$, on S, and the radiation condition (7). From the uniqueness theorem for the exterior electric boundary value problem [6] we obtain $\text{curl}\, u_p = 0$ in D_e.

That $u := u_p + u_s$ is a solution of $\Delta^* u + \kappa_s^2 u = 0$ in D_e, follows by straightforward calculations. Since the cartesian components of every solution of the vector Helmholtz equation satisfying the radiation condition (7) also satisfy the radiation condition of Sommerfeld [6, see Corollary 4.14], we obtain that u fulfills the radiation condition (3).

That u fulfills the boundary conditions (2) is easily seen by

$$\begin{aligned} u &= u_s + u_p = n(n, u_s + u_p) - [n, [n, u_s + u_p]] \\ &= n(n, f) - [n, [n, f]] = f, \text{ on } S. \end{aligned}$$

From the uniqueness theorem for Problem D we obtain the following uniqueness

Theorem 3.2: Problem HD possesses at most one solution if for κ_p and κ_s the condition (8) holds.

Proof: Let u_p, u_s be a solution of the homogeneous Problem HD. As in the proof of Theorem 3.1 we can see that $u := u_s + u_p$ is a solution of Problem D but now to the homogeneous boundary condition. Therefore, by the uniqueness theorem for the exterior displacement problem we derive $u = 0$ in D_e.

Now we have $u_s = -u_p$ in D_e and there holds $\text{div}\, u_p = 0$ and $\text{curl}\, u_s = 0$ in D_e. From this we conclude

$$-\kappa_j^2 u_j = \Delta u_j = \text{grad div}\, u_j - \text{curl curl}\, u_j = 0,$$

and therefore $u_j = 0$ in D_e, for $j = s, p$. This means that Problem HD possesses at most one solution.

Submitted on September 22, 2011 / Accepted on September 30, 2011

References

1. Müller C. Foundations of the Mathematical Theory of Electromagnetic Waves. Springer-Verlag, Berlin, Heidelberg, New York, 1969.
2. Kupradze V. D. Potential Methods in the Theory of Elasticity. Israel Program for Scientific Translations, Jerusalem, 1965.
3. Ahner J. A. The exterior time harmonic elasticity problem with prescribed displacement vector on the boundary. *Archiv der Mathematik*, 1976, v. XXVII, 106–111.
4. Hähner P., Hsiao G. C. Uniqueness theorems in inverse obstacle scattering of elastic waves. *Inverse Problems*, 1993, v. 9, 525–534.
5. Leis R. Vorlesungen über partielle Differentialgleichungen zweiter Ordnung. Bibliographisches Institut, Mannheim, 1967.
6. Colton D., Kress R. Integral Equation Methods in Scattering Theory. John Wiley & Sons, New York, 1983.

On the Neutrino Opera in the CNGS Beam

Armando V.D.B. Assis
Departamento de Física, Universidade Federal de Santa Catarina – UFSC, Trindade 88040–900, Florianópolis, SC, Brazil.
armando.assis@pgfsc.ufsc.br

In this brief paper, we solve the relativistic kinematics related to the intersection between a relativistic beam of particles (neutrinos, e.g.) and consecutive detectors. The gravitational effects are neglected, but the effect of the Earth rotation is taken into consideration under a simple approach in which we consider two instantaneous inertial reference frames in relation to the fixed stars: an instantaneous inertial frame of reference having got the instantaneous velocity of rotation (about the Earth axis of rotation) of the Cern at one side, the lab system of reference in which the beam propagates, and another instantaneous inertial system of reference having got the instantaneous velocity of rotation of the detectors at Gran Sasso at the other side, this latter being the system of reference of the detectors. Einstein's relativity theory provides a velocity of intersection between the beam and the detectors greater than the velocity of light in the empty space as derived in this paper, in virtue of the Earth rotation. We provide a simple calculation for the discrepancy between a correct measure for the experiment and a measure arising due to the effect derived in this paper.

1 Definitions and Assumptions

Consider a position vector for CERN in relation to the center of the Earth, vector $\vec{C}$, and a position vector for the Gran Sasso receptors in relation to the center of the earth, vector $\vec{G}$. Consider the angular velocity vector of the Earth along its axis of rotation, vector $\vec{\omega}$. The velocity of rotation of $\vec{C}$ in relation to Earth's axis is given by $\vec{v}_C = \vec{\omega} \times \vec{C}$. Analogously, the velocity of rotation of $\vec{G}$ in relation to Earth's axis is given by $\vec{v}_G = \vec{\omega} \times \vec{G}$. Consider a baseline $\mathcal{L}_{CG}$ connecting $\vec{C}$ and $\vec{G}$ along the vector $\vec{G} - \vec{C}$; CERN's and Gran Sasso's latitudes $(\updownarrow_S^N)$, λ_C and λ_G, respectively, and CERN's and Gran Sasso's longitudes $(\leftarrow WE \rightarrow)$, α_C and α_G, respectively.

Since the effect related to the velocity of the neutrinos depends on its own velocity at the completion of the calculation and on the rotation of the Earth, viz., such effect does not depend on the specific values of the lateral velocity (to be defined below) of the receptors, as we will see, we may consider some geometric assumptions to simplify the geometry related to the baseline path $\vec{G} - \vec{C}$ along $\mathcal{L}_{CG}$ through the Earth.

Firstly, we will consider $\vec{C}$ and $\vec{G}$ having got the same latitude λ^*:

$$\lambda_C = \lambda_G = \lambda. \tag{1}$$

These latitudes would be important if the effect to be derived here was related to specific values of latitude, its fluctuations, systematic and/or statistical errors related to it etc., related to the six standard deviations that characterizes the claim related to the experiment. But that is not the case. Now, consider the plane $\mathbf{\Pi}$, orthogonal to $\vec{\omega}$, that cross the Earth through the hypothetically common latitude containing the points $\vec{C}$ and $\vec{G}$. Trace two lines pertaining to $\mathbf{\Pi}$: the line $\mathcal{L}_{CA}$, from the point $\vec{C}$ to Earth's rotation axis, and the line $\mathcal{L}_{GA}$, from the point $\vec{G}$ to Earth's rotation axis. $\mathcal{L}_{CA}$ and $\mathcal{L}_{GA}$ cross the rotation axis at the point $\vec{A}$. Also, trace the mediatriz line $\mathcal{L}_{MA}$, from the point $\vec{A}$ to the point $\vec{M} = (1/2)\left(\vec{G} + \vec{C}\right)$, equally dividing $\mathcal{L}_{CG}$. The angle between $\mathcal{L}_{CA}$ and $\mathcal{L}_{MA}$ equals the angle between $\mathcal{L}_{MA}$ and $\mathcal{L}_{GA}$, being this angle given by:

$$\alpha = \frac{1}{2}\left(\alpha_G - \alpha_C\right). \tag{2}$$

Upon the previous remarks regarding the geometric simplifications, the same remarks hold for the radius of the Earth, i.e., we will consider the Earth as a sphere. Thus, the following relation holds:

$$\left|\vec{C}\right| = \left|\vec{G}\right| = R_E = \frac{R}{\cos\lambda}, \tag{3}$$

where R_E is the radius of the Earth, its averaged value $R_E = 6.37 \times 10^6 m$, and $R = \left|\vec{C} - \vec{A}\right| = \left|\vec{G} - \vec{A}\right|$.

2 Defining two Instantaneous Inertial Reference Frames

The relativistic kinematics will run in the plane $\mathbf{\Pi}$ previously defined. The line $\mathcal{L}_{CG}$ will define an axis: Ox, with the origin O at the point $\vec{C}$, being the unitary vector of the axis Ox, $\hat{e}_x$, given by:

$$\hat{e}_x = \frac{\vec{G} - \vec{C}}{\left|\vec{G} - \vec{C}\right|}. \tag{4}$$

Now, define the Oz axis contained in the $\mathbf{\Pi}$ plane such that its unitary vector, $\hat{e}_z$, is given by:

$$\hat{e}_z = -\hat{e}_x \times \frac{\vec{\omega}}{\left|\vec{\omega}\right|}. \tag{5}$$

*The latitudes of CERN and Gran Sasso are, respectively: $46^{\text{deg}}14^{\text{min}}3^{\text{sec}}$(N) and $42^{\text{deg}}28^{\text{min}}12^{\text{sec}}$(N). The longitudes of CERN and Gran Sasso are, respectively: $6^{\text{deg}}3^{\text{min}}19^{\text{sec}}$(E) and $13^{\text{deg}}33^{\text{min}}0^{\text{sec}}$(E).

The axis Oy is trivially obtained with its unitary vector being given by:

$$\hat{e}_y = \hat{e}_z \times \hat{e}_x. \quad (6)$$

Now, we define the system at the Gran Sasso's detectors, $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$, such that its origin $\tilde{O}$ is at the point $\vec{G}$, being the unitary vector of the axis $\tilde{O}\tilde{x}$, the same $\hat{e}_x \equiv \hat{e}_{\tilde{x}}$. The axis $\tilde{O}\tilde{z}$ is parallel to Oz, with the same unitary vector $\hat{e}_z \equiv \hat{e}_{\tilde{z}}$, with analogous reasoning to obtain the axis $\tilde{O}\tilde{y}$ and its unitary vector $\hat{e}_{\tilde{y}} \equiv \hat{e}_y$. In other words, $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$ is the parallel pure translation of $Oxyz$ from $\vec{C}$ (CERN) to $\vec{G}$ (Gran Sasso).

To define the two instantaneous inertial reference frames to accomplish, simply, the effect of the Earth rotation, we, firstly, write down the rotation velocities of the points $\vec{C}$ and $\vec{G}$ about Earth's rotation axis, i.e., we write down the rotation velocities of (CERN) and (Gran Sasso) about Earth's axis. For CERN, the rotation velocity $\vec{v}_C$ reads:

$$\vec{v}_C = \vec{\omega} \times \vec{C} = \omega R \hat{e}_\phi = \omega R\,(\cos\alpha\,\hat{e}_x - \sin\alpha\,\hat{e}_z)\,, \quad (7)$$

where the auxiliar unitary vector has been the azimutal ϕ-versor of the spherical coordinates, the spherical coordinates with origin at the center of the Earth with its equatorial dextrogyre plane Ξ such that $\vec{\omega} \cdot \vec{\xi} = \vec{0}\ \forall\ \vec{\xi} \in \Xi$. For Gran Sasso, the rotation velocity $\vec{v}_G$ reads:

$$\vec{v}_G = \vec{\omega} \times \vec{G} = \omega R \hat{e}_{\tilde{\phi}} = \omega R\,(\cos\alpha \hat{e}_{\tilde{x}} + \sin\alpha\,\hat{e}_{\tilde{z}})\,, \quad (8)$$

where $\hat{e}_{\tilde{\phi}}$ is the azimutal ϕ-versor previously defined, but now at Gran Sasso.

We see via the eqs. (7) and (8) that both the frames of reference, $Oxyz$ and $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$, are instantaneously under a null relative translation through the common axis $Ox \equiv \tilde{O}\tilde{x}$, and under a reverse translation through their respective parallel axes $Oz \parallel \tilde{O}\tilde{z}$. We will *inertially* consider this quite instantaneous* effect of the reverse translation (Newton's first law will hold, we will instantaneously neglect the gravitational field through the neutrino travel to Gran Sasso, as well as the weak characteristic for neutrino interactions with matter) via the following approach:

- We will consider a system of reference $O_C x_C y_C z_C$ that exactly coincides with $Oxyz$ at the instant $t_C = 0$, but with the following constant velocity of translation in relation to the fixed stars: $\vec{v}_\epsilon = \omega R \cos\alpha\,\hat{e}_{x_C} - \omega R \sin\alpha\,\hat{e}_{z_C} = \omega R \cos\alpha\,\hat{e}_x - \omega R \sin\alpha\,\hat{e}_z = \vec{v}_C$, such that the neutrino travel will be in this inertial referential. The subscript ϵ is to asseverate this referential is being considered for the neutrino travel during the entire process (emission→detection), but with $\epsilon \approx 0$ in the sense given in the previous footnote [$\epsilon \equiv \delta t_v / T << 1$]. Considered this, we will drop the subscript C in $O_C x_C y_C z_C$, for the sake of economy of notation, and rename it simply as $Oxyz$, although this latter is not the original one;
- We will consider a system of reference $O_G x_G y_G z_G$ that exactly coincides with $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$ at the instant $t_G = t_C = 0^\dagger$, but with the following velocity of translation in relation to the fixed stars: $\vec{u}_\epsilon = \omega R \cos\alpha\,\hat{e}_{x_G} + \omega R \sin\alpha\,\hat{e}_{z_G} = \omega R \cos\alpha\,\hat{e}_{\tilde{x}} + \omega R \sin\alpha\,\hat{e}_{\tilde{z}} = \vec{v}_G$. Considered this, we will drop the subscript G in $O_G x_G y_G z_G$, for the sake of economy of notation, and rename it simply as $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$, although this latter is not the original one;
- We will consider a system of reference travelling with the beam of neutrinos, but this will be explained in the next section.

3 From CERN to the Flux through the Gran Sasso Detectors

From now on, we model the lattice (strips, emulsion, cintilators etc) distribution through the Grand Sassos' detectors from the perspective of an $Oxyz^\ddagger$ observer with the following characteristics:

- The average proper (no Lorentz contraction in $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$) displacement of detectors along $\tilde{O}\tilde{x}$ is $d_{\tilde{O}\tilde{x}}$;
- The average proper (no Lorentz contraction in $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$) displacement of detectors along $-\tilde{O}\tilde{y}$ is $d_{\tilde{O}\tilde{y}}$;
- The average proper (no Lorentz contraction in $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$) displacement of detectors along $-\tilde{O}\tilde{z}$ is $d_{\tilde{O}\tilde{z}}$;
- The detectors in $Oxyz$ will be abstracted to a tridimensional $d_{0x} \times d_{0y} \times d_{0z}$ othogonally spaced lattice falling upward [see the eqs. (7) and (8)] at the velocity $\vec{v}_G - \vec{v}_C = 2\omega R \sin\alpha\,\hat{e}_z$, being the basis vectors of these sites given by $\{\vec{d}_{0x} = d_{0x}\hat{e}_x,\ \vec{d}_{0y} = -d_{0y}\hat{e}_y,\ \vec{d}_{0z} = -d_{0z}\hat{e}_z\}$, where $\{\hat{e}_x,\ \hat{e}_y,\ \hat{e}_z\}$ is the canonical spacelike 3D euclidian orthonormal basis of $Oxyz$.
- We will neglect relativistic (Einstein's) effects related to the movement of the lattice of detectors, the movement of $\tilde{O}\tilde{x}\tilde{y}\tilde{z}$ in $Oxyz$, as previously stated, but such effects will become important in the referential of the neutrino beam (to be defined below).

Now, we define the neutrino frame of reference $O'x'y'z'$ in the canonical configuration with the frame of reference $Oxyz$, i.e., coincident origins at $t = t' = 0$ keeping the spacelike parallelism of the axes $x \equiv x'$, $y \equiv y'$ and $z \equiv z'$ and

*The time spent by a neutrino beam to accomplish the race from $\vec{C}$ to $\vec{G}$, δt_v, obey $\delta t_v / T << 1$, where T is the period of Earths's rotation about its axis, thus quite instantaneous in relation to the Earth daily kinematics.

†The relativistic effects between the systems of reference at CERN and at Gran Sasso related to time synchronization is being neglected due to the order of magnitude related to the velocities due to the Earth rotation and due to the magnitude of the gravitational field as previously stated. Furthermore, we are undressing these effects between these systems at $\vec{C}$ and $\vec{G}$ to asseverate the relevant relativistic effects that will lead to the neutrino velocity will raise in virtue of relativistic motion in relation to the detectors in Gran Sasso, as we will see.

‡From now on, we are working with the inertial frames (in relation to the fixed stars) defined above, viz., from now on: $Oxyzt$ means $O_C x_C y_C z_C t_C$ (see the two final paragraphs of the previous section); $\tilde{O}\tilde{x}\tilde{y}\tilde{z}\tilde{t}$ means $O_G x_G y_G z_G$ (see the two final paragraphs of the previous section).

boosted with velocity $v_\nu\hat{e}_x$ in relation to the $Oxyz$ frame. Consider a neutrino beam entering the block of detectors in Gran Sasso in the $Oxyz$ frame of reference. The beam passes a lattice of detectors stated above, being these detectors raining upward with velocity $\vec{v}_G - \vec{v}_C = 2\omega R\sin\alpha\hat{e}_z$ through the beam in the $Oxyz$ frame of reference. A horizontal neutrino beam, along Ox, may contact a horizontal [vertical means along Oz, upward means in the $\hat{e}_z$ direction, and horizontal means parallel to the xy plane] lattice of detectors parallelly raining upward in virtue of the Earth rotation as discussed before (raining upward with velocity $\vec{v}_G - \vec{v}_C = 2\omega R\sin\alpha\hat{e}_z$). Once an interaction occurs between the horizontal beam and consecutively located detectors in this horizontal lattice, this interaction is simultaneous in the $Oxyz$ (rigorously $Oxyzt$, but the context is clear here) world, implying non-simultaneity for these raindrops of detectors in the $O'x'y'z'$ world. The distribution of these raindrops of detectors must have, *instantaneously* at t' in $0'x'y'z'$ world, the following characteristics:

- The displacement between two consecutive raindrops of detectors correlated to the respective simultaneous ones in $Oxyz$, these latter displaced by the proper distance $x_{i+1} - x_i = d_{0x}$ along Ox and belonging to the falling upward xy plane of detectors in $Oxyz$, is given by:

$$x'_{i+1}(t') - x'_i(t') = \gamma^{-1}(x_{i+1} - x_i) = \gamma^{-1}d_{0x}, \tag{9}$$

being $\gamma = 1/\sqrt{1 - v_\nu^2/c^2}$, c the speed of light in the empty space, v_ν the speed of the neutrino, whose velocity is along the $\hat{e}_x$ direction in the $Oxyz$ world (the beam of neutrinos is at rest in its referential $O'x'y'z'$, as previously seem).

- The displacement between two consecutive raindrops of detectors correlated to the respective simultaneous ones in $Oxyz$, these latter displaced by the proper distance $z_i - z_{i+1} = 0$ along Oz and belonging to the falling upward xy plane of raining detectors in $Oxyz$, is given by:

$$z'_{i+i}(t') - z'_i(t') = 2\frac{v_\nu d_{0x}}{c^2}\omega R\sin\alpha. \tag{10}$$

- The vertical distance between consecutive (consecutive but inclined in the $O'z'y'z'$ world; the parallel to xy planes of detectors parallelly raining upward in $Oxyz$ become inclined in $O'x'y'z'$) raining planes of detectors Π'_i and Π'_{i+1}, $\forall\, i$, remains the same d_{0z} distance, the distance between consecutive parallelly raining planes of detectors. The raining upward planes turn out to be inclined in relation to the $x'y'$ plane of the neutrino world $O'x'y'z'$ by the angle:

$$\theta = \pi - \arctan\left(2\frac{\gamma v_\nu}{c^2}\omega R\sin\alpha\right). \tag{11}$$

Indeed, let's derive these facts. Firstly, instantaneously at t in $Oxyz$, two consecutive raindrops* Ox along, are time delayed in $O'x'y'z'\times\{t'\}$ world by the amount:

$$t'_{i+1} - t'_i = \gamma\left(t - \frac{v_\nu}{c^2}x_{i+1}\right) - \gamma\left(t - \frac{v_\nu}{c^2}x_i\right) = -\gamma\frac{v_\nu}{c^2}(x_{i+1} - x_i) = -\gamma\frac{v_\nu}{c^2}d_{0x}, \tag{12}$$

in virtue of the Lorentz transformations $(x,t)\to(x',t')$. Here, we see a detection that occurs at the position x_{i+1} pertaining to the horizontal lattice of detectors in Gran Sasso, at the plane $\tilde{x}\tilde{y}$ within the block of detectors in Gran Sasso, hence more internal, (remember $\tilde{x}\tilde{y}\parallel xy$), *must occur earlier* than the detection at the position x_i in the frame of reference of the beam of neutrinos, and the i-raindrop is late in relation to the $(i+1)$-raindrop. Hence, backwarding the t'_i clocks down to the the t'_{i+1} instant (backwarding the movie, maybe better: backwarding the neutrino's opera), i.e., comparing the non-simultaneous events in the beam of neutrinos frame, the event $i+1$ ocurring when the $i+1$-raindrop crosses the beam of neutrinos and the event i when the i-raindrop crosses the beam of neutrinos (remember these events are simultaneous in $Oxyz$) previously to infer the instantaneous (at t'_{i+1}) position of the i-raindrop when the $i+1$ raindrop crosses the beam of neutrinos at the instant $t'_{i+1} < t'_i$ in the $O'x'y'z't'$ frame, the i-raindrop must move the amounts (backwarding the movie from the instant t'_i at which the i-raindrop crosses the beam of neutrinos in the $O'x'y'z't'$ world to the non-simultaneous instant $t'_{i+1} < t'_i$ at which the $i+1$-raindrop crosses the beam of neutrinos in the $O'x'y'z't'$ world): $\delta z'$ downward and $\delta x'$ to the right, being these amounts given by:

$$\delta z' = \left(\frac{2\omega R\sin\alpha}{\gamma}\right)\times\left(-\gamma\frac{v_\nu}{c^2}d_{0x}\right) = -\frac{2\omega d_{0x}v_\nu R\sin\alpha}{c^2};$$

$$\delta x' = (-v_\nu)\times\left(-\gamma\frac{v_\nu}{c^2}d_{0x}\right) = \frac{v_\nu^2\gamma d_{0x}}{c^2}, \tag{13}$$

since $(-v_\nu\hat{e}_{x'} + (2\omega R\sin\alpha/\gamma)\hat{e}_{z'})$ is the velocity of raindrops in $0'x'y'z'$, obtained from the Lorentz transformations $L(\vec{u})$ for the 3-velocities of the Gran Sasso lattice block of sensors, the raining raindrops lattice of sensors, from the $Oxyz$ to the beam of neutrinos frame $O'x'y'z'$:

$$(0,0,2\omega R\sin\alpha)|_{Oxyz}\xrightarrow{L(\vec{u})}(-v_\nu,0,2\omega R\sin\alpha/\gamma)|_{O'x'y'z'}\,. \tag{14}$$

But, at t, the i-raindrop and the $(i+1)$-raindrop have got the same z coordinate, since they are in a xy plane, and, since the $z\to z'$ Lorentz map is identity, these raindrops must have the same z' coordinate at their respective transformed instants

*From now on, we will call raindrops the detectors in the lattice of detectors within the block of detectors at Gran Sasso. Thus, raindrops ≡ detectors within the lattice of detectors defined at the beginning of this section; 1 raindrop ≡ 1 detector within the lattice of detectors within the block of detectors at Gran Sasso.

(of course, since at *each* of these transformed instants, different instants in $O'x'y'z'$ in virtue of the non-simultaneity in this frame, the z' coordinate will read the same, since these reaindrops will cross the beam and the beam has the same coordinate z' in its own frame of reference, viz., the beam is parallel to $O'x'$). Hence, backwarding t'_i clocks down to the the t'_{i+1} instant, one concludes that the $\delta z'$ in the eq. (13) is the instantaneous, at same t', height shift between consecutive raindrops that simultaneously cross the beam of neutrinos in $Oxyz$. The $x \to x'$ Lorentz map is not identity, implying one must calculate the $x'_{i+1} - x'_i$ shift at the $Oxyz$ instantaneous t:

$$x'_{i+1}(t) - x'_i(t) = \gamma(x_{i+1} - v_\nu t) - \gamma(x_i - v_\nu t) = \gamma(x_{i+1} - x_i)$$

$$= \gamma d_{0x}. \tag{15}$$

This shift is related to different instants, t'_i, t'_{i+1}, in the beam of neutrinos frame. Thus, backwarding t'_i clocks down to the the t'_{i+1} instant (backwarding the movie to observe the earlier t'_{i+1} instantaneous), this amount given by the eq. (15) is reduced by the amount $\delta x'$ given by eq. (13):

$$x'_{i+1}(t') - x'_i(t') = \gamma d_{0x} - \gamma d_{0x}\frac{v_\nu^2}{c^2} = \gamma d_{0x}\left(1 - \frac{v_\nu^2}{c^2}\right)$$

$$= \gamma^{-1} d_{0x}. \tag{16}$$

The first eq. (13) gives the eq. (10), since eq. (13) gives the z' position of the i-raindrop at the previous instant t'_{i+1} before the i-raindrop crosses the beam of neutrinos in the $O'x'y'z'$, therefore:

$$z'_i\left(t'_{i+1}\right) = z'_\nu - \frac{2\omega d_{0x} v_\nu R\sin\alpha}{c^2}, \tag{17}$$

where z'_ν is *a* constant z' coordinate of the beam of neutrinos in its own frame; and, since the z' position of the $(i+1)$-raindrop at the t'_{i+1} instant is z'_ν (due to the very fact the $(i+1)$-raindrop crosses the beam at the instant t'_{i+1} in the $O'x'y'z'$ world), one has $z'_{i+1}\left(t'_{i+1}\right) = z'_\nu$, from which, with the eq. (17), one has got:

$$z'_{i+1}\left(t'_{i+1}\right) - z'_i\left(t'_{i+1}\right) = 2\frac{v_\nu d_{0x}}{c^2}\omega R\sin\alpha, \tag{18}$$

reaching the eq. (10). The non-instantaneous displacement (non-instantaneous in $O'x'y'z'$) given by eq. (15) is the distance between two sucessive non-instantaneous interactions with the beam, raindrops marks assigned upon the beam in $O'x'y'z'$. This fact is easy to understand, as these instantaneously assigned marks (instantaneous in $Oxyz$) would become splayed in $O'x'y'z'$, since the beam turns out to be contracted in $Oxyz$ due to Lorentz contraction. Also, one shall infer that eq. (16) gives the t' instantaneous displacement of falling upward raindrops along $O'x'$. The reason why the distance between consecutive raindrops marks γd_{0x} are bigger than the contracted distance $\gamma^{-1}d_{0x}$ of the two consecutive falling raindrops is explained by the non-simultaneity between these raindrops when touching the proper beam in the $0'x'y'z'$ world, straightforwardly seem by the inclination (the horizontal planes of raindrops in $Oxyz$ inclines in $O'x'y'z'$) between the raindrop plane containing these two consecutives raindrops in $O'x'y'z'$ and the proper plane $\Pi'_\nu \parallel x'y'$ containing the neutrinos beam in $O'x'y'z'$; i.e., when the first sensor raindrop crosses the beam, assigning the first interaction, the second travels an amount $\delta x'$ to the left given by the second eq. (13) before crossing the beam, assigning the second interaction. A xy instantaneous falling upward plane of sensors within the block of sensors at Gran Sasso containing raindrops in $Oxyz$ world becomes an inclined instantaneous falling upward plane in $O'x'y'z'$ world, being the inclination, eq. (11), easily derived from eqs. (16) and (18):

$$\tan(\pi - \theta) = \frac{\delta z'(t')}{x'_{i+1}(t') - x'_i(t')} = 2\frac{\gamma v_\nu}{c^2}\omega R\sin\alpha, \tag{19}$$

giving the eq. (11).

4 Faster than Light Effects in Gran Sasso

To understand the effect, first, consider two sensors, say i-raindrop and $(i+1)$-raindrop. If these sensors are constructed to tag the instants, t_{i+1} and t_i, at which two events are registered at their exact locations and a team of physicists obtains the time variation interval by $t_{i+1} - t_i$, being $x_{i+1} - x_i$ the distance between these sensors, one would have:

$$\frac{\delta x}{\delta t} = \frac{x_{i+1} - x_i}{t_{i+1} - t_i} = \infty, \tag{20}$$

for simultaneous events ($t_i = t_{i+1}$), if one expects a signal is travelling between the sensors. Furthermore, if one expects a privileged direction along which the signal should travel from the i-raindrop (first) to the $(i+1)$ raindrop (later), if the $(i+1)$-raindrop registered a signal before the i-raindrop, violating the expected sequential direction of detections, one would say the signal would have been registered from the future to the past direction. In the previous section the instantaneous events in the $Oxyz$ became non-instantaneous in the beam frame of reference, and the internal register within the Gran Sasso block along the direction $Ox \equiv \tilde{O}\tilde{x}$, at the position x_{i+1} registered the interaction with the beam at the same instant the internal register at the position x_i registered, since these events were hypothetically simultaneous in $Oxyz$, in virtue of the Earth rotation. From the point of view of the neutrino beam, these registers occurred in the order: x'_{i+i} before, x_i later, due to the inclination of the raindrops planes in virtue of the Earth rotation. We are forced to conclude the rotation of the Earth may provide a kinematics of intersection between beams and sequential sensors that may led to the conclusion the sensors are registering time intervals related to quasi-simultaneous events that are cintilated by different particles at different positions almost at the same time, leading to an errouneous conclusion that the signal would have travelled between the sensors generating the time tag data. E.g.,

suppose two ideal clocks, perfect ones, gedanken ones, that register the instants: t_i at which a beam of neutrinos enters the block of raindrop sensors in Gran Sasso and t_o at which this beam of neutrinos emerges from the block. Let d_B be the lenght travelled through the block. One team of physicists will measure the velocity of the beam by $d_B/(t_o - t_i)$ with no use of data from the sensors within the block. Another team will perform the calculation from the data obtained from a sequence of sensors (raindrops) located Ox along. This second team may obtain registers at diferent positions x_{i+1} and x_i related to the lateral intersection between these sensors and the beam entirely into the block of sensors but with the beam travel not entirely accomplished through the block. The data of this second team would be mistaken, since the registers at the different locations x_{i+1} and x_i would not have been made by the same neutrino, implying the clocks at x_{i+1} and x_i would be registering two quasi-simultaneous events *not related to a same neutrino*, concluding erroneously that the time variation between these events was so small that the particle that generated these events would be travessing with a velocity greater than c.

Einstein's theory of relativity does not avoid velocities greater than the light in the empty space, but avoids *an unique* particle propagating with velocity greater than the velocity of light in the empty space. To infer that a velocity greater than c may arise from the discussion through this brief article, consider the velocity two *different* raindrops interact with the beam of neutrinos in the beam $O'x'y'z'$ frame of reference. These events are non-simultaneous in the beam frame as previously discussed, but the beam crosses two sucessive interactions with a propagation that is faster than c, since the distance between two sucessive interactions along the beam in the beam frame of reference is given by the eq. (15), γd_{0x}, being the time spent given by the eq. (12), $(\gamma v_\nu d_{0x})/c^2$. Thus, the 2-propagation $\mathcal{V}'_{(i+1)\to(i)}$ (the number 2 to denote two bodies are related to a single propagation velocity):

$$\mathcal{V}'_{(i+1)\to(i)} = \frac{c}{v_\nu}c \Rightarrow \mathcal{V}'_{(i+1)\to(i)} > c. \tag{21}$$

As asseverated this is not a propagation of a single particle, but a ratio between the covered distance along the beam in the beam frame and the time interval spent to interact, non-simultaneously, with two sequential but distinct sensors (raindrops x_{i+1} and x_i). Of course, if $v_\nu \to 0$, these distinct interactions will tend to become simultaneous, leading to the result discussed at the beginning of this section (eq. 20). It follows that is not difficult to conclude that the time elapsed between two distinct sensors must be related to just an unique particle if one is intended to use their time tags for velocity computations.

5 The Consequence of the Effect

A simple calculation provides the discrepancy obtained by the set of CNGS detections intended to obtain the correct velocity of the neutrino particle announced few days ago. Let v_c be the correct value for the neutrino's velocity, δS_1, the distance between the CERN and the point at which the neutrino enters the block of detectors at Gran Sasso, δS_2 the lenght of the block of detectors to be internally covered by the neutrino at Gran Sasso, δt_1 the elapsed time spent by the neutrino to cover δS_1, δt_2 the elapsed time spent by the neutrino to cover δS_2, δt_f a fake elapsed time due to the effect previously discussed, and v_f a fake velocity that would arise from an erroneous measure for the elapsed time through δS_2. Hence:

$$v_c = \frac{\delta S_1 + \delta S_2}{\delta t_1 + \delta t_2}, \tag{22}$$

$$v_f = \frac{\delta S_1 + \delta S_2}{\delta t_1 + \delta t_f}. \tag{23}$$

A simple calculation gives:

$$\frac{v_f}{v_c} = 1 + \frac{\delta t_2 - \delta t_f}{\delta t_1 + \delta t_f} \Rightarrow \frac{v_f - v_c}{v_c} = \frac{\delta t_2 - \delta t_f}{\delta t_1 + \delta t_f}. \tag{24}$$

One should not write $v_f = \delta S_2/\delta t_f$, since, as previously discussed, δt_f is not related to a propagation of a particle, but to the time elapsed between two detectors in the same xy plane in the $Oxyz$ frame of reference. If a sequence of cintilations within the block of detectors are generated by different neutrinos due to the effect previously discussed, and this sequence is interpreted as a path traced by a single neutrino, the measure of the distance covered within the block of detectors at Gran Sasso would encapsulate an error for each estimated path having got the effect encrusted within it. This distance, an erroneous one, is δS_2, although this distance may be defined as a correct one for purposes of comparison with a case in which (hypothetically) a neutrino travelled this distance with the correct velocity v_c. Obviously, the δt_f would not be related to the time spent to cover this distance, since, in a case in which the effect, as previously explained, was generated due to simultaneous time tagging at two different sensors (raindrops) due to two different neutrinos in the $Oxyz$ reference frame, one would have $\delta t_f = 0$ for a continuous laterally traced path in virtue of the Earth rotation, from which $\delta S_2 \neq 0$. δS_2 would arise, under the effect discussed in this paper, from a path misinterpretation. But, once one defines δS_2 as the distance to be covered in a comparison case with a neutrino used to cover it with the correct velocity v_c:

$$\delta t_2 = \frac{\delta S_2}{v_c}, \tag{25}$$

turns out to be the correct elapsed time. From the eq. (24), one reaches:

$$v_f\left(\delta t_1 + \delta t_f\right) - v_c\delta t_1 = v_c\delta t_2. \tag{26}$$

If the computation is done taken into consideration the erroneous *elapsed* time δt_f, being this $\delta t_f = 0$ in a case of

simultaneity due to the effect previously discussed, the eq. (26) reads:

$$v_f \delta t_1 - v_c \delta t_1 = v_c \delta t_2, \tag{27}$$

and the effect turns out to increase the fake velocity in relation to the correct velocity in a manner in which the neutrino with the fake velocity would be winning the race by an amount of distance given by $\delta S_2 = v_c \delta t_2$, accomplished the path δS_1, from the eq. (27). Hence, eq. (24) turns out to read:

$$\frac{v_f - v_c}{v_c} = \frac{\delta t_2}{\delta t_1} = \frac{\delta S_2}{v_c} \frac{v_c}{\delta S_1} = \frac{\delta S_2}{\delta S_1}. \tag{28}$$

With the values [1]:

$$\frac{v_f - v_c}{v_c} = \left[2.48 \pm 0.28\ (stat.) \pm 0.30\ (sys.)\right] \times 10^{-5}, \tag{29}$$

and:

$$\delta S_1 = 733 \times 10^3 m, \tag{30}$$

we reach for the discrepancy between the covered distances:

$$\delta S_2 = \left[18.2 \pm 2.05\ (stat.) \pm 2.20\ (sys.)\right] m. \tag{31}$$

6 Conclusion

We conclude the relativistic effect discussed here in virtue of the Earth rotation may lead to a misinterpretation of the elapsed time within the block of detectors at Gran Sasso.

Acknowledgments

A.V.D.B.A is grateful to Y.H.V.H and CNPq for financial support.

Submitted on September 27, 2011 / Accepted on October 2, 2011

References

1. The OPERA collaboration: T. Adam et al. Measurement of the neutrino velocity with the OPERA detector in the CNGS beam http://arxiv.org/abs/1109.4897 *arXiv:1109.4897*, 2011.

Open Letter by the Editor-in-Chief: Declaration of Academic Freedom (Scientific Human Rights)
The Portuguese Translation[*]

Declaração de Liberdade Acadêmica

(Direitos Humanos Científicos)

Artigo 1: Preâmbulo

O início do século XXI reflete, mais do que em qualquer outra época na história da Humanidade, a profundidade e a significância do papel da Ciência e da tecnologia nas relações humanas.

A natureza poderosamente pervasiva da Ciência e da tecnologia modernas tem levado a uma percepção corriqueira de que descobertas-chave somente podem ser feitas, ou principalmente, por intermédio de grandes grupos de pesquisa sob tutela corporativa ou governamental com acesso à instrumentação de alto custo e a vasto suporte pessoal.

Todavia, a percepção corriqueira é mítica, esconde a verdadeira natureza de como as descobertas científicas são realmente feitas. Enormes e caros projetos tecnológicos, complexos ou não, não são senão o resultado da aplicação das profundas percepções de pequenos grupos de pesquisadores dedicados ou cientistas solitários, frequentemente trabalhando isolados. Um cientista trabalhando sozinho está e estará, tanto agora quanto no futuro, assim como ocorrera no passado, apto a fazer uma descoberta que pode influenciar substancialmente o destino da humanidade e mudar a face de todo o planeta sobre o qual nós tão insignificantemente residimos.

Descobertas revolucionárias são geralmente feitas por indivíduos trabalhando em posições subordinadas dentro de agências governamentais, em instituições de ensino e pesquisa, ou em empresas comerciais. Consequentemente, o pesquisador, com frequencia, está vinculado ou limitado por diretores de instituições e corporações que, trabalhando em uma prioridade diferente, visam controlar e aplicar perquisa e descoberta científicas para benefício pessoal, organizacional, ou engrandecimento pessoal.

O registro histórico de descobertas científicas está repleto de instâncias de supressão e ridicularização por parte do poder estabelecido, já há muito se tendo revelado e reivindicado pela inexorável marcha de necessidade prática e iluminação intelectual. Também assim se tem corrompido e sujado o registro histórico por plágio e deliberada perversão de fatos, perpetrados pelos inescrupulosos, motivados por inveja e avareza. Assim também o é hoje em dia.

O objetivo desta declaração é manter e incentivar a doutrina fundamental de que a pesquisa científica deve estar livre de latentes e abertas influências repressivas advindas de diretivas burocráticas, políticas, religiosas e pecuniárias, e de que a criação científica é um direito humano, não menos do que outros de tais direitos e árduas esperanças que se propuserem em tratados e leis internacionais.

Todos os cientistas que a apoiam devem ser fiéis a esta Declaração, como uma indicação de solidariedade para com a comunidade científica internacional interessada, e para conceder o Direito dos cidadãos do mundo à livre criação científica de acordo com suas habilidades e disposição individuais, para o avanço científico, por sua extrema habilidade como cidadãos decentes em um mundo indecente, [para] o benefício da Humanidade.

Artigo 2: Quem é um cientista

Um cientista é qualquer pessoa que faz Ciência. Qualquer pessoa que colabora com um cientista no desenvolvimento e proposição de idéias e dados em pesquisa ou aplicação é também um cientista. A posse de uma qualificação formal não é um pré-requisito para que uma pessoa seja um cientista.

Artigo 3: Onde a Ciência é produzida

A pesquisa científica pode ser desenvolvida em qualquer lugar, por exemplo, em um lugar de trabalho, durante um curso formal de educação, durante um programa acadêmico patrocinado, em grupos, ou, de modo independente, por indivíduos em suas casas.

Artigo 4: Liberdade de escolha do tema de pesquisa

Muitos cientistas que trabalham por graus de pesquisa mais avançados ou em outros programas de pesquisa em instituições acadêmicas tais como universidades e centros de estudos avançados são privados de trabalhar em um tema de pesquisa de sua própria escolha por acadêmicos seniores e/ou funcionários administrativos, não por falta de instrumentos de apoio, mas, em vez disso, por causa de hierarquia acadêmica e/ou pelo que outros funcionários não aprovam a linha de pensamento em virtude de seu potencial conflito com dogma estabelecido, teorias favorecidas, ou financiamento de projetos outros que possam ser desacreditados pela pesquisa proposta. A autoridade da maioria ortodoxa é muito frequentemente invocada para percalçar um projeto de pesquisa tal que a autoridade e seus pressupostos não sejam incomodados. Essa prática comum é uma obstrução deliberada ao livre pensamento científico, não sendo científica ao extremo, e criminosa. Ela não pode ser tolerada.

[*]Original text published in English: *Progress in Physics*, 2006, v. 1, 57–60. Online — http://www.ptep-online.com/

Versão original em Língua Inglesa por Dmitri Rabounski, editor-chefe do periódico *Progress in Physics*. E-mail: rabounski@ptep-online.com.

Traduzido para a Língua Portuguesa por Armando V. D. B. Assis. E-mail: armando.assis@pgfsc.ufsc.br.

Um cientista trabalhando para qualquer instituição acadêmica, autoridade ou agência, deve ser completamente livre para escolher um tema de pesquisa, limitado apenas pela disponibilidade de recursos materiais e aptidões intelectuais aptas a serem oferecidas pela instituição educacional, agência ou autoridade. Se um cientista procede com a pesquisa sendo membro de um grupo colaborativo, os diretores de pesquisa e os líderes de equipe devem estar limitados ao contexto de colsultores e conselheiros em relação ao que fora escolhido por um cientista do grupo como sendo um tema de pesquisa relevante.

Artigo 5: Liberdade de escolha de métodos de pesquisa

Frequentemente, tem-se o caso de se exercer pressão sobre um cientista por parte de funcionários administrativos ou acadêmicos seniores em relação a um programa de pesquisa conduzido dentro de um ambiente acadêmico, de se forçar um cientista a adotar métodos de pesquisa alheios aos que o cientista escolheu, por nenhuma razão outra senão preferência pessoal, viés, política institucional, mando editorial, ou autoridade coletiva. Essa prática, a qual está muito espalhada, é uma negação deliberada de liberdade de pensamento e não pode ser permitida.

Um cientista não comercial ou acadêmico tem o direito de desenvolver um tema de pesquisa em qualquer caminho razoável e por quaisquer meios que considere ser os mais efetivos. A decisão final sobre como a pesquisa será conduzida deve ser feita pelo próprio cientista.

Caso um cientista não comercial ou acadêmico trabalhe como membro de uma equipe colaborativa não comercial ou acadêmica de cientistas, os líderes de projeto e diretores de pesquisa deverão ter apenas direitos de conselheiros e consultores, não devendo, todavia, influenciar, mitigar ou limitar os métodos de pesquisa ou tema de pesquisa de um cientista dentro do grupo.

Artigo 6: Liberdade de participação e colaboração em pesquisa

Há um elemento significativo de rivalidade institucional na prática de Ciência moderna, concomitante com elementos de inveja pessoal e preservação de reputação a todo custo, independente das realidades científicas. Isso tem, com frequencia, levado os cientistas a não convidar colegas competentes de instituições rivais ou outros sem afiliação acadêmica. Essa prática também é uma obstrução deliberada ao progresso da Ciência.

Caso um cientista não comercial ou acadêmico requeira a colaboração de outrem e este estiver de acordo em oferecê-la, aquele cientista terá a liberdade de convidar este outro para que lhe preste essa ou qualquer outra ajuda, supondo-se que a assistência requerida esteja sob um contexto pressuposto de pesquisa associada. Se tal colaboração estiver fora desse escopo pressuposto, o cientista ainda terá a liberdade de escolher o outro para sua discrição, livre de qualquer interferência de quem quer que seja.

Artigo 7: Liberdade de discordar em discussão científica

Em consequencia de ciúmes furtivos e interesse adquirido, a Ciência moderna repudia discussões abertas e premeditadamente bane aqueles cientistas que questionam as visões ortodoxas. Muito frequentemente, cientistas de habilidade extraordinária, que apontam deficiências em teorias vigentes ou em interpretação de dados, são rotulados de crackpots (excêntricos tolos), de modo às suas interpretações serem convenientemente ignoradas. Eles são particular e privadamente ridicularizados e são barrados dos congressos científicos, seminários e colóquios de modo às suas idéias não encontrarem audiência. Falsificação deliberada de dados e má representação de teoria são ferramentas frequentes e atuais dos inescrupulosos na supressão de fatos, tanto técnicos quanto históricos. Comitês internacionais de meliantes científicos têm sido formados e esses conselhos abrigam e dirigem convenções internacionais nas quais apenas os seus acólitos são permitidos a apresentar artigos, independente da qualidade de conteúdo. Esses conselhos angariam grades somas de dinheiro público para financiar seus projetos patrocinados, por intermédio de engano e mentira. Qualquer objeção às suas propostas, ainda que objetadas sobre bases científicas, é silenciada por quaisquer meios disponíveis, de modo que o dinheiro continue fluindo para dentro de suas contas de projeto, e que os garanta empregos bem pagos. Cientistas que se opõem a essa praxe têm sido exonerados a pedido daqueles; outros têm sido impedidos de ocupar posições acadêmicas por uma rede de cúmplices corruptos. Em outras situações, alguns têm sido expulsos de suas candidaturas a programas de educação superior tal como doutoramento, por expressar idéias que minam uma teoria da moda, a despeito do arraigo temporal que uma teoria ortodoxa todavia possa ter. O fato fundamental de que nenhuma teoria científica é definitiva e inviolável, estando portanto aberta à discussão e à reexaminação, é completamente ignorado. O fato de que um fenômeno pode ter um número de explicações plausíveis é também ignorado, e maliciosamente põem em descrédito qualquer explicação que não esteja de acordo com a opinião ortodoxa, recorrendo, sem arguir exceções, ao uso de argumentos não científicos para justificar suas opiniões tendenciosas.

Todos os cientistas devem ser livres para discutir sua pesquisa e a pesquisa de outrem sem medo de que sejam pública ou privadamente ridicularizados sem fundamento, de que sejam acusados, depreciados, impugnados ou postos em descrédito de qualquer outra forma por alegações não consubstanciadas. Nenhum cientista deve ser posto em uma posição pela qual seu sustento ou reputação estejam sob risco em consequencia de expressão de uma opinião científica. A liberdade de expressão científica deve ser suprema. O uso de autoridade na refutação de um argumento científico não é científico e não deve ser usada para amordaçar, suprimir, intimidar, ostraci-

zar, ou, por qualquer forma coercitiva, barrar um cientista. A supressão deliberada de fatos científicos ou argumentos, seja por ato ou omissão, e a manipulação deliberada de dados para sustentar um argumento ou para por em descrédito uma visão oposta é fraude científica, perfazendo um crime científico. Princípios de evidência devem guiar toda discussão científica, seja tal evidência física, teórica ou também uma combinação.

Artigo 8: Liberdade para publicar resultados científicos

Uma censura deplorável de artigos científicos tem agora se tornado a prática padrão dos conselhos editoriais dos maiores jornais e arquivos eletrônicos, e de seus bandos de alegados árbitros especialistas. Os árbitros são em sua maior parte protegidos pelo anonimato de modo que um autor não tem como verificar suas alegadas especialidades. Artigos são atual e rotineiramente rejeitados caso o autor não concorde com ou contradiga uma teoria preferida ou a corrente ortodoxa principal. Muitos artigos são agora automaticamente rejeitados em virtude do aparecimento na lista de artigos de autor de um cientista em particular que não tenha encontrado favor entre os editores, entre os árbitros, ou entre outros censores especialistas, sem qualquer consideração que seja feita ao conteúdo do artigo. Existe uma listagem negra de cientistas dissidentes e esta lista é comunicada aos e entre os participantes de conselhos editoriais. Tudo isso contribui para o aumento da não isenção, da tendenciosidade, e para a punível supressão de livre pensamento, devendo ser condenado pela comunidade científica internacional.

Todos os cientistas devem ter o direito de apresentar seus resultados de pesquisa científica, no todo ou em parte, em conferências científicas relevantes, de publicar os mesmos em jornais científicos impressos, em arquivos eletrônicos, e em qualquer outro meio. Nenhum cientista deve ter seus artigos ou relatórios rejeitados quando submetidos à publicação em jornais científicos, em arquivos eletrônicos, ou em outro meio, simplesmente por que seu trabalho questiona a opinião majoritária corrente, por que conflita com as visões de um conselho editorial, pelo que mine as bases de outra corrente ou projetos de pesquisa planejados por outros cientistas, por estar em conflito com qualquer dogma político, credo religioso, ou opinião pessoal de outrem; e nenhum cientista deve ser inserto em listagem negra ou por outra forma censurado, impedido, por quem quer que seja, de publicar. Nenhum cientista deve bloquear, modificar, ou de outra forma interferir na publicação de um trabalho de cientista sob promessa de presentes ou qualquer forma de suborno.

Artigo 9: Coautoria de artigos científicos

É um segredo alardeado em círculos científicos, que muitos coautores de artigos de pesquisa têm, de fato, pouca ou nenhuma relação com as pesquisas ali relatadas. Muitos supervisores de estudantes graduados, por exemplo, não são contrários à colocação de seus nomes em artigos escritos por aquelas pessoas que apenas nominalmente trabalham sob suas supervisões. Em muitos de tais casos, a pessoa que de fato escreve o artigo tem um intelecto superior ao do supervisor nominal. Em outras situações, novamente pelos propósitos de notoriedade, reputação, dinheiro, prestígio, e os similares, as pessoas não participantes são incluídas no artigo como coautoras. Os autores de fato de tais artigos podem apenas objetar sob risco de serem subsequentemente penalizados de alguma forma, ou mesmo de serem expulsos de suas candidaturas a graduações de pesquisa mais elevadas ou de suas equipes de pesquisa, conforme o caso. Muitas têm de fato sido expulsas sob tais circunstâncias. Essa prática horrorosa não pode ser tolerada. Apenas aquelas pessoas responsáveis pela pesquisa deveriam ter autoria oficialmente reconhecida.

Nenhum cientista deve convidar outra pessoa para ser incluída e nenhum cientista deve permitir que seu nome seja incluso como coautor de um artigo científico caso não tenha contribuído significativamente para a pesquisa que se relata no artigo. Nenhum cientista deve permitir que ele mesmo ou ela mesma seja coagido por qualquer representante de uma instituição acadêmica, corporação, agência governamental, ou qualquer outra pessoa, a incluir seu nome como coautor em relação à pesquisa para qual não tenha significativamente contribuído, e nenhum cientista deve permitir que seu nome seja usado como coautor à guisa de presentes ou quaisquer subornos. Nenhuma pessoa deve induzir ou tentar induzir um cientista, por qualquer forma, a um caminho que permita que o nome do cientista seja incluído como coautor de um artigo científico relacionado a assuntos para os quais não tenha significativamente contribuído.

Artigo 10: Independência de afiliação

Muitos cientistas estão agora empregados sob contratos de curto período. Com o término do contrato de emprego, também termina a afiliação acadêmica. É frequente a política de conselhos editoriais em que pessoas sem uma afiliação acadêmica ou comercial não publiquem. Na ausência de afiliação, muitos recursos não estão disponíveis ao cientista, e as oportunidades de apresentar palestras e artigos em conferências são reduzidas. Essa é uma prática viciosa que deve ser parada. Ciência não reconhece filiação.

Nenhum cientista deve ser impedido de apresentar artigos em conferências, colóquios ou seminários, de publicar em qualquer meio, de acessar bibliotecas acadêmicas ou publicações científicas, de participar de encontros científicos, ou de dar conferências, por necessitar de afiliação a uma instituição acadêmica, instituto científico, laboratório governamental ou comercial, ou de qualquer outra organização.

Artigo 11: Acesso aberto à informação científica

Muitos livros especializados sobre assuntos científicos e muitos jornais científicos rendem pouco ou nenhum lucro, de modo que editores comerciais não estão dispostos a publicá-

los sem uma contribuição em dinheiro de instituições acadêmicas, agências governamentais, fundações filantrópicas, e correlatos. Sob tais circunstâncias, editores comerciais deveriam permitir acesso livre a versões eletrônicas das publicações, e esforçar-se por manter o custo dos materiais impressos num mínimo.

Todos os cientistas deverão se esforçar para assegurar que seus artigos de pesquisa estejam disponíveis à comunidade científica internacional de modo gratuito, ou alternativemante, se tal não puder ser evitado, a um mínimo custo. Todos os cientistas deverão tomar medidas ativas para fazer com que seus livros técnicos estejam disponíveis ao custo mais baixo possível de modo a poder estar a informação disponível à mais ampla comunidade científica internacional.

Artigo 12: Responsabilidade ética de cientistas

A História testifica que descobertas científicas são usadas tanto para fins bénéficos quanto malévolos, para o benefício de alguns e para a destruição de outros. Dado que o progresso da Ciência e da tecnologia não pode parar, meios para que se contivesse a aplicação malévola deveriam ser estabelecidos. Apenas um governo democraticamente eleito, laico, com liberdade racial e não tendencioso, pode salvaguardar a civilização. Apenas governos, tribunais e comitês democraticamente eleitos podem salvaguardar o direito de livre criação científica. Hoje em dia, vários estados não democráticos e regimes totalitários conduzem pesquisa ativa em física nuclear, química, virologia, engenharia genética etc, com propósito de produzir armas nucleares, químicas e biológicas. Nenhum cientista deveria ter interesse em colaborar com estados não democráticos ou regimes totalitários. Qualquer cientista coagido a trabalhar no desenvolvimento de armas para tais estados deveria encontrar meios de diminuir o progresso de programas de pesquisa e de reduzir a produção científica de forma que a civilização e a democracia possam finalmente prevalecer.

Todos os cientistas adquirem uma responsabilidade moral por suas criações científicas e descobertas. Nenhum cientista deve voluntariamente se engajar no desenho ou construção de armas de qualquer tipo, para o que quer que seja, para estados não democráticos ou regimes totalitários ou permitir que suas habilidades científicas e conhecimentos sejam aplicados no desenvolvimento do que quer que seja prejudicial à Humanidade. Um cientista deve viver pelo dito de que todo governo não democrático e toda violação de direitos humanos são crime.

22 de novembro de 2005 / Traduzido em 17 de maio de 2011

Open Letter by the Editor-in-Chief: Declaration of Academic Freedom (Scientific Human Rights)
The Arabic Translation*

إعلان الحرية الأكاديمية
(الحقوق العلمية للأنسانية)

المادة ١ : مقدمة

إن بداية القرن الحادى و العشرون بعد الميلاد تعكس أكثر من أي وقت مضى في تاريخ البشرية عمق وأهمية دور العلم والتكنولوجيا للبشرية. ونظرا لطبيعة الأنتشار بقوة المميزة للعلم والتكنولوجيا الحديثين فأنه يمكن إحراز اكتشافات ثورية أخرى سواء كانت بمشاركة او بشكل حصرى من قبل الحكومة أو المجموعات البحثية المؤسسية الممولة مما يؤهلها للوصول إلى أجهزة باهظة الثمن بشكل كبير وجحافل من الباحثين و الأفراد.

إن الإدراك و الأتفاق المشترك للعلم هو امر من المستحيل ويناقض الطبيعة الحقيقية لكيفية صنع الاكتشافات العلمية. ذلك لأن المشاريع التكنولوجية الكبيرة والمكلفة – مهما كانت معقدة – ليست سوى نتيجة لتطبيق الأفكار العلمية العميقة لمجموعات صغيرة من الباحثين أو العلماء المنفردين الذين يعملون – في كثير من الأحيان – في عزلة. إن عالم واحد – سواء في الحاضر والمستقبل – وحده هو تماما كما في الماضي قادرة اكتشاف ما يمكن أن يؤثر بشكل كبير على مصير الإنسانية وتغيير وجه هذا الكوكب الذي نقطنه.

تتم عادة الاكتشافات الرائدة من قبل أفراد يعملون في مناصب ثانوية داخل الوكالات الحكومية أومؤسسات البحث العلمى أوالتعليمي أو المؤسسات التجارية. وبالتالي فإن الباحث في كثير من الأحيان مقيد أو يتم قمعه من قبل مديري المؤسسات والشركات الذين يعملون لأهداف مختلفة، ساعون إلى السيطرة وتطبيق الاكتشافات العلمية والبحوث من أجل الربح الشخصي أو المؤسسي أو حتى الترقي الشخصي.

إن السجل التاريخي لاكتشاف العلمي مليء بحالات القمع والسخرية من قبل المؤسسات و تنكشف هذه الحالات و تبرر بعدها بسنوات بسبب المسير الغير قابل للتغيير للعلم و الضرورة العملية والتنوير الفكري. كذلك حالات الإشقاء وتلطيخ السمعة و الانتحال والتحريف المتعمد و التي يرتكبها عديمي الضمير- بدافع الحسد والجشع- وهكذا هو الحال حتى الان.

الهدف من هذا الإعلان هو دعم وتعزيز مبادئ البحث العلمي الأساسية التي يجب أن تكون خالية من التأثيرات القمعية العلنية والمستترة و من التوجيهات والبيروقراطية السياسية والمالية (إن عدم الأعتقاد فى توجيهات الأديان هو خروج عن اخلاق العلم وعدم أتباع لحق الأنسان في الاعتقاد بالله و قد يقصد هنا التوجيهات الزائدة عن منهج الدين، المترجم) والمالية، وأن صنع العلم هو حق من حقوق الإنسان لا تقل عن تلك الحقوق و الأمال الأخرى كما نصت عليه المواثيق الدولية والقانون الدولي.

يجب دعوة جميع العلماء المناصرين بالألتزم بهذا الإعلان باعتباره مؤشرا للتضامن مع المجتمع العلمي الدولي المعني وتعاطفا مع حقوق المواطنة ذلك من أجل صنع إنجازات علمية دون قيود وفقا لمهاراتهم و تصوراتهم الفردية - من أجل النهوض بالعلم – وذلك لأقصى قدرتهم كمواطنين شرفاء فى تلك البيئة الغير ملائمة. تلك المنفعة هى العائد على البشرية بعدما امضت العلوم والتكنولوجيا فترة طويلة جدا في مطال الاضطهاد.

المادة ٢ : من هو العالم

العالم هو أي شخص يتناول العلم و أي شخص يتعاون مع أحد العلماء في تطوير والإقتراح الأفكار والبيانات في مجال البحث العلمي أو تطبيقه هو أيضا عالما. و أن اكتساب مؤهل رسمي ليس شرطا مسبقا للإنسان لكى يكون عالما.

المادة ٣ : ما الذى ينتج العلم

ويمكن إجراء البحث العلمي في أي مكان على الإطلاق، على سبيل المثال في مكان العمل أو خلال مقرر دراسى أو خلال برنامج اكاديمي مدعوم وذلك خلال مجموعات أو كأفراد في اطار إجراءات مستقلة.

المادة ٤ : حرية اختيار اطار البحث

مُنع العديد من العلماء الذين يعملون للحصول على درجة أعلى في برامج البحث العلمى أو غيرها من برامج البحوث في المؤسسات الأكاديمية مثل الجامعات والكليات من العمل على موضوع بحث من اختيارهم وذلك من قبل المشرفين الأكاديميين أو المسؤولين الإداريين، وليس ذلك لعدم وجود امكانيات فنية لكن - و بدلا من ذلك - بسبب التسلسل الهرمي الأكاديمي و/ أو غيره من الأمور الرسمية و ذلك ببساطة لعدم أتفاقه و العقيدة السائدة أو النظريات المفضلة أو مسار التمويل للمشاريع الأخرى التي قد تفقدت مصداقيتها نتيجة البحوث المقترحة. وكثيرا ما ثارت السلطة الغالبة لافشال مشروع بحثي بحيث لا يزعج السلطة والميزانيات. هذه الممارسة الشائعة هي عرقلة متعمدة للفكر العلمي الحر و هي غير علمية في نهايتها وتعد جنائية و بالتالي لا يمكن التغاضي عنها.

أن اي عالم يعمل في أي مؤسسة أكاديمية أو هيئة أو وكالة يكون حرا تماما في اختيار موضوع البحث، مرتبطا فقط بالدعم المادي والمهارات الفكرية القادرة على أكتسابها من المؤسسات التعليمية أو الهيئات أو الوكالات. إن قام أحد العلماء بإجراء البحوث بوصفه عضوا في مجموعة تعاونية فإنه يقتصر على مديريه أو قادة الفريق البحثى القيام بالأدوار الاستشارية والتشاور في ما يتعلق باختيار موضوع البحوث ذات الصلة.

المادة ٥ : حرية اختيار طرق البحث

يتكرر كثيرا أن يتم التسبب بإجبار العالم من قبل أفراد إداريين أو كبار الأكاديميين في ما يتعلق ببرنامج البحوث التي أجريت في البيئة الأكاديمية على أتباع أساليب بحثية مختلفة عن تلك التي اختارها العالم

*Original text published in English: *Progress in Physics*, 2006, v.1, 57–60. Online — http://www.ptep-online.com/

E-mail: rabounski@ptep-online.com. Progress in Physics النص الأصلي كتب بواسطة ديمترى رابونسكى رئيس تحرير مجلة
Email: emaghraby@techemail.com. مصر Elsayed K. Elmaghraby و ترجم للعربية بواسطة

إن من حق العالم الغير هادف للربح أو الغير أكاديمي تطوير موضوع البحث بأي وسيلة و بأي سبب معقول يعتبره أكثر فعالية. القرار النهائي بشأن الكيفية التي سيتم بها إجراء البحوث سيقوم بها العالم وحده.
إذا كان عالما غير هادف للربح أو غير أكاديمي يعمل كعضو في فريق تعاوني من العلماء غير هادف للربح أو غير أكاديمي فإن قادة ومدراء المشاريع البحثية لديهم حقوق النصح و الأرشاد فقط ويجب ألا يؤثروا خلاف ذلك بتخفيف أو تقييد طرق البحث أو البحوث المعنية لعالم داخل المجموعة.

المادة ٦ : حرية المشاركة والتعاون في مجال البحوث

هناك عنصر للتنافس بين المؤسسات في ممارسة العلم الحديث - بالتزامن مع عناصر من الحسد الشخصي والحفاظ على السمعة مهما كان الثمن - بغض النظر عن الحقائق العلمية. وقد أدى ذلك في كثير من الأحيان إلى منع العلماء من التعاون بين الزملاء المختصصين في المؤسسات المتنافسة أو غيرهم من الذين دون أي انتماءات أكاديمية. هذه الممارسة هي أيضا عرقلة متعمدة لتقدم العلمي.
إذا طلب عالما غير هادف للربح مساعدة من شخص أخر بعد موافقته فإن للعالم الحرية في دعوة هذا الشخص لتقديم أي مساعدة بشرط تضمين المساعدة في ميزانية البحوث المرتبطة بها. إذا كانت المساعدة مستقلة عن أي اعتبارات مالية فإن للعالم وحده الحرية في الانخراط في طلب المساعدة من الشخص و في تقديره بشكل خالي من أي تدخل على الإطلاق من قبل أي شخص مهما كان.

المادة ٧ : حرية الاختلاف في المناقشة العلمية

بسبب الغيرة الدفينة وتشعب الأهتمامات المكتسب يتجنب العاملون في العلم الحديث المناقشة مفتوحة عمدا ويتجنب أولئك العلماء الذين يشككون من وجهات النظر المتعصبة. في كثير من الأحيان يوصف العلماء ذوي القدرة البارزة - الذين يشيرون إلى أوجه القصور في النظرية الحالية أو الى تفسير بياناتها بالمجازيب حتى أنه يحدث تجاهل لوجهات نظرهم بشكل عفوي ويكونون - سرا وعلانية - وبصورة منهجية ممنوعون من حضور المناقشات العلمية والحلقات الدراسية والندوات بحيث اصبحت أفكارهم لا تجد طريقة للاعلان عن نفسها. إن التزييف المتعمد للبيانات والسخرية من نظرية ما هي الآن أدوات معتادة من عديمي الضمير في طمس الحقائق، سواء الفنية أوالتاريخية. وقد تم تشكيل لجان دولية مِن من يوصفون بالعلماء واستضافة هذه اللجان ومباشرتها للمقابلات الدولية التي يسمح بها فقط لمساعديهم بتقديم أوراق بحثية بغض النظر عن جودة المحتوى. هذه اللجان استقطبت مبالغ كبيرة من المال من الأموال العامة لرعاية تمويل مشاريعها من خلال اللجوء إلى الخداع والكذب. وأي اعتراض على مقترحاتهم على أسس علمية يتم منعه بأي وسيلة تحت تصرفهم، لذلك يمكن ان يستمر تدفق الاموال الى حسابات مشايعهم التي تضمن لهم وظائف جيدة الأجر. وقد تم عزل العلماء المعارضين لهم بأمر منهم و منع الآخرين من الحصول على التعيينات الأكاديمية من خلال شبكة من المتواطئين. في حالات أخرى تم طرد البعض من الترشيح في برامج الدرجة الأعلى مثل الدكتوراه للتعبير هم عن الأفكار التي تقوض "نظرية المألوف". مع ذلك هذه النظرية موجودة منذ فترة طويلة و قد تكون تقليدا دينيا عمره الاف السنين. الحقيقة الأساسية بأنه لا توجد نظرية علمية محددة وغير قابلة للانتهاك - وبالتالي فهي مفتوحة للمناقشة وإعادة النظر – تم تجاهلها تماما. كأن نتجاهل حقيقة أن أي ظاهرة قد يكون لها عدد من التفسيرات المعقولة - وهذا لا يتفق مع الرأي التقليدي-و تستخدم و دون تردد الحجج الغير علمية لتبرير آراء متحيزة.
جميع العلماء يجب أن يكونوا احرارا في مناقشة أبحاثهم وبحوث الآخرين دون خوف من السخرية العامة و خاصة تلك التي لا تملك أساسا ماديا أو محض اتهام أو استخفاف أو طعنا فيه أو خلاف ذلك مما لا أساس له من ادعاءات فقدت مصداقيتها. و لا يجب وضع اي عالم في موقف يعرض اسلوب معيشتة أو سمعته للخطر بسبب التعبير عن الرأي العلمي. حرية التعبير عن الرأي يجب أن تكون الأولى بالتقدير. استخدام السلطة في دحض حجة علمية ليست علما ويجب ألا تستخدم لإسكات أوقمع أوترهيب أو نبذ أو غير ذلك من إكراه أو تقييد للعالم. القمع المتعمد للحقائق العلمية أو الحجج إما بفعل أو الامتناع عن فعل والجبر المتعمد للبيانات لدعم حجة أو لتشويه سمعة الرأي المخالف والغش العلمي الأمور التي تصل إلى حد الجريمة العلمية. ومبادئ الأستدلال يجب أن تحرك المناقشة العلمية و يكون الدليل المادي أو النظري أو مزيج منها هما الأساس.

المادة ٨ : حرية نشر النتائج العلمية

الآن أصبحت الرقابة ضعيفة على الأوراق العلمية و من الممارسات العامة المعمول بها في هيئات تحرير عدد من المجلات الرئيسية ومكتبات الحفظ الإلكترونية وعلى شرائح من محكمي المجلات المزعومين. إن المحكم هو القسم الاكبر المحمي بهدف عدم الكشف عن الهويته حتى أن العالم الذي يتقدم للنشر لا يمكن التحقق من خبراته المفترضة. والآن و بشكل روتيني يتم رفض البحوث إذا لم يتفق المؤلف أو تعارض مع النظرية المفضلة لدى المحكم أوالعقيدة السائدة. و في الوقت الحالي يتم رفض العديد من الأوراق البحثية تلقائيا بشكل ظاهري بتأثير ظهور العالم في قائمة مؤلفات عالم اخر ليس مقبولا لدى احد من المحررين أوالمحكمين أو حتى المرقبين و الخبراء الأخرين دون أي اعتبار على الإطلاق لمحتويات الورقة البحثية. هناك قوائم سوداء للعلماء المعارضين وتبلغ هذه القائمة بين هيئات التحرير للمجلات المتعاونة. كل هذه التراكم من الانحياز والقمع للتفكير الحر وجب أدانتها من قبل المجتمع العلمي الدولي.
جميع العلماء لديهم الحق في تقديم نتائج بحوثهم العلمية، كليا أو جزئيا، في المؤتمرات العلمية ذات الصلة، ونشرها في المجلات العلمية المطبوعة والالكترونية أو بأية وسيلة أخرى. لا يجوز رفض الأوراق العلمية او التقارير لأي عالم عند تقديمها للنشر في المجلات العلمية المطبوعة والإلكترونية أو/ وسائل الإعلام الأخرى وذلك– و بكل بساطة – بسبب نقاط بحوثهم المعنية أو تعارضها مع وجهات النظر لهيئة التحرير أو تقويضها للأسس الحالية أو الأسس المخطط لها لمشاريع البحوث من قبل العلماء الآخرين أو أن تتعارض مع أي عقيدة سياسية أو ملية (مع الأخذ في الاعتبار العامل الاخلاقي للبحث الاكاديمي، المترجم) أو رأي شخصي آخر. ولا يجب وضع أي عالم في القوائم السوداء أو لومه على خلاف ذلك ومنعه من النشر من قبل أي شخص أي كان. لا يجوز لأي عالم حجب أو تعديل أو التداخل في نشر بحث بتقديم الوعود أو أي مظهر اخر من مظاهر الرشوة من أي نوع.

المادة ٩ : المشاركة في تأليف المنشورات العلمية
من الصعب كتمان أنه في كثير من الأروقة العلمية تكون مشاركة المؤلفين فى الأوراق البحثية في الواقع ضئيلة أو لا علاقة لهم بتاتا بالبحث الذى ذكروا فيه. فالعديد من المشرفين على طلاب الدراسات العليا- على سبيل المثال - لا يتورعوا عن وضع أسمائهم على أوراق مكتوبة من قبل أولئك الأشخاص الذين وضعوا اسميا تحت إشرافهم. في حالات كثيرة من هذه يكون للشخص الذي يكتب الورقة في الواقع سبق التفكير على المشرف الاسمي و في حالات أخرى – و لأغراض الشهرة والسمعة والمال والمكانة وغيره مما سبق ذكره- لا يشاركون نهائيا فى الورقة أو البحث الذى تضمن اسمائهم. يمكن للمؤلفين الفعلي لهذه الأوراق الأعتراض مما يعرضهم لخطر العقاب لاحقا بطريقة ما أو حتى أستبعاده من التقدم للترشيح لدرجة علمية أو ادارية أعلى فى الفريق البحثي، كما هو الحال. وقد تم بالفعل طرد الكثير في مثل هذه الظروف. لا يمكن السكوت عن هذه الممارسة المرعبة. يجب فقط على الأشخاص المسؤولين عن البحث أعتماد مصداقية المشاركة فيه.
لا يجوز لأي عالم أن يدعو شخص آخر إلى إدراج أسمه ولا يجوز تضمين اسماء باعتبارهم مؤلفون مشاركون في ورقة علمية اذا لم يسهموا إسهاما كبيرا في الدراسة التى نشرت. لا يجوز لأي عالم أو عالمة السماح بإكراه نفسه او نفسها - من قبل ممثل أي مؤسسة أكاديمية أو تعاونية أو وكالة حكومية أو أي شخص آخر- بأن تدرج أسمائهم كمؤلف مشترك للبحوث التي لم يساهم فيها بشكل كبير ولا يجوز لأي عالم السماح بأستخدام اسمهم كمؤلف مشارك في مقابل أي هدايا أو أى رشاوى أخرى. لا يجوز لأي شخص التحريض أو محاولة التحريض لإقناع العالم بأى طريقة كانت للسماح بتضمين اسم باحث ومؤلف كمشارك في ورقة علمية تتعلق بالمواضيع لم يسهم فيها الاخير بشكل ملحوظ.

المادة ١٠ : الاستقلال عن الانتماء
يعمل الآن كثير من العلماء في إطار عقود قصيرة الأجل و مع انتهاء عقد العمل ينتهي الانتماء الأكاديمي لهذا العالم. و غالبا ما تكون سياسة هيئات التحرير أن لا تنشر بحوثا للأشخاص بدون انتماء أكاديمي أو تجاري. في غياب الانتماء فأن كثيرة من الموارد ليست متاحة للعالم ويتم تقليص فرصه لتقديم ورقات البحث في المناقشات والمؤتمرات. هذا الممارسة مفرغة و يجب أن تتوقف فالعلم لا يعترف بالانتماء. لا يجوز منع أي عالم من تقديم الأوراق البحثية في المؤتمرات أوالندوات أو الحلقات الدراسية أونشرها في أي من وسائل الإعلام، لا يجوز منع أي عالم من الوصول إلى المكتبات الأكاديمية أو المنشورات العلمية ولا يجوز منعه من حضور الاجتماعات العلمية أو من إعطاء المحاضرات لعدم وجود انتماء مع المؤسسة الأكاديمية أو المعاهد العلمية أوالحكومة أو المختبرات التجارية أو أي منظمة أخرى.

المادة ١١ : الحصول على المعلومات العلمية
كثير من الكتب المتخصصة في الأمور العلمية وكذلك المجلات العلمية تدر أرباح ضئيلة أو معدومة بحيث يكون الناشرين التجاريين غير راغبين في نشرها بدون مساهمة مالية من المؤسسات الأكاديمية أوالهيئات الحكومية أوالمؤسسات الخيرية وما شابه ذلك. في ظل هذه الظروف ينبغي على الناشرين التجاريين السماح بحرية الوصول إلى النسخ الالكترونية من المنشورات و السعى جاهدين للحفاظ على تكلفة المواد المطبوعة إلى الحد الأدنى.

على جميع العلماء الجهاد لضمان أن الأبحاث الخاصة بهم متاحة للمجتمع العلمي الدولي مجانا أو بدلا من ذلك - إذا كان لا يمكن تجنب عدم المجانية - أن تكون بأقل تكلفة. كما ينبغي على جميع العلماء اتخاذ تدابير فعالة لجعل كتبهم التقنية متاحة بأقل تكلفة ممكنة بحيث يمكن للمعلومات العلمية أن تكون متاحة على اوسع نطاق للمجتمع العلمي الدولي.

المادة ١٢ : المسؤولية الأخلاقية للعلماء
يشهد التاريخ ان الاكتشافات العلمية تستخدم لاغراض الخير و اغراض الشر على نحو سواء وأيضا لمنفعة البعض وضرر للآخرين. ولأن تقدم العلم والتكنولوجيا لا يمكن أن يتوقف فإن بعض وسائل احتواء التطبيقات المضرة ينبغي أن تنشأ. كما أن الحكومات المنتخبة ديمقراطيا فقط والخالية من التحيز الديني والعرقي وغيرها هى التى يمكنها الحفاظ على الحضارة فإن الإدارات المنتخبة ديمقراطيا ولجان التحكيم فقط يمكنها حماية حق حرية صناعة العلم. و اليوم دول مختلفة ربما غير ديمقراطية أو أنظمة شمولية تقوم بإجراء البحوث النشطة في الفيزياء النووية والكيمياء وعلم الفيروسات والهندسة الوراثية وما إلى ذلك من أجل إنتاج أسلحة نووية وكيماوية وبيولوجية. لا ينبغي لعالم أن يتعاون طوعا مع الدول غير الديمقراطية أو النظم الشمولية. وينبغي على أي عالم أُكره للعمل على تطوير الأسلحة إيجاد طرق أو وسائل ليبطئ التقدم المحرز في برامج البحوث والإنتاج العلمي للحد منه لكى تكون الحضارة والديمقراطية هى السائدة في نهاية المطاف.
يتحمل كل العلماء المسؤولية الأخلاقية لإبداعاتهم العلمية واكتشافاتهم. لا يجوز لأي عالم الانخراط طوعا في تصميم أو بناء أسلحة من أي نوع على الإطلاق وذلك لدول غير ديمقراطية أو الأنظمة الشمولية أو السماح بتطبيق المهارات العلمية له أو لها والمعرفة لتطوير أي شيء ضارة بالبشرية. للعالم أن يعيش تحت شعار أن كل حكومة غير ديمقراطية وانتهاك لحقوق الإنسان هو جريمة.

كتبه : ديمتري رابونسكى ،
رئيس تحرير مجلة Progress in Physics فى ٢٢ نوفمبر ٢٠٠٥

LETTERS TO PROGRESS IN PHYSICS

Scientists Deduced the Existence of Particles with Faster-than-Light Speeds Recently Discovered by CERN

Ion Pătraşcu
"Fraţii Buzeşti" National College, Craiova, Romania
E-mail: patrascu_ion@yahoo.com

In this paper we present a short survey on Smarandache Hypothesis that there is no speed barrier in the universe and one can construct arbitrary speeds, hypothesis which has been partially confirmed by the recent CERN results of OPERA team led by Antonio Ereditato that experimentally found that neutrino particles travel faster than *c*.

Physicists at CERN have recently experimentally discovered particles traveling faster than light: the neutrinos! The OPERA experiment, which sent sprays of neutrinos from CERN, Geneva, Switzerland, to INFN Gran Sasso Laboratory, Italy, found that neutrinos traveled underground 730 km faster than light could do. Dr. Antonio Ereditato of University of Bern, leader of the OPERA scientist team, made the results public and invited scientists all over the world to discuss these astonishing results.

There are mediums where the light travels slower than some particles, for example in water and oil, but not in the vacuum. There also exist superluminal phenomena like wave phase velocity and wave group velocity, but in these cases no information or energy travels faster than the light. Similarly are the X-waves whose superluminal velocity of the peak is a transitory phenomenon, but their wavefronts move with the speed *c* [1].

In the breaking News on September 22, 2011, in the Live Science.com, it is said that proven true, the laws of physics have to be re-written [2].

Professor Florentin Smarandache from the University of New Mexico, United States, has deduced the existence of particles moving faster-than-light in a published paper called "There Is No Speed Barrier in the Universe" in 1998 [3], as an extension of a 1972 manuscript [4] that he also presented in 1993 at the conference *Paradoxism in Literature and Science* held in the Universidad de Blumenau, Brazil. His paper is based on the Einstein-Podolsky-Rosen Paradox [5], a Bohm's paper [6], and Bell's Inequalities [7]. For this goal known as *Smarandache Hypothesis*, and for his neutrosophic logic, set, and probability (which are the most general and powerful logic, and, respectively, set and probability theories today), Prof. Smarandache awarded the Telesio-Galilei Academy Gold Medal in 2010. Smarandache Hypothesis is also included in the Weinstein's *Encyclopedia of Physics* [8]. It is is enounced as follows:

- Suppose a certain physical process produces a pair of entangled particles A and B (having opposite or complementary characteristics), which fly off into space in the opposite direction and, when they are billions of miles apart, one measures particle A; because B is the opposite, the act of measuring A instantaneously tells B what to be; therefore those instructions would somehow have to traveled between A and B faster than the speed of light; hence, one can extend the Einstein-Podolsky-Rosen paradox and Bell's inequalities and assert that the light speed is not a speed barrier in the universe;
- Even more, one can construct any speed, even greater than the speed of light, *c*, by measuring particle A at various time intervals;
- Also, the information from particles A and B is transmitted instantaneously (thus, there is no speed barrier in the universe).

Although superluminal phenomena are in contradiction with Einstein's theory of Special Relativity (1905) that prevents energy, information and (real) mass from traveling faster than light, Smarandache (1972) considered that superluminal phenomena do not violate Causality Principle, neither produce time traveling, nor necessitating infinite energy for particles traveling at speeds greater than the speed of light.

Submitted on September 26, 2011 / Accepted on October 03, 2011

References

1. Maiorino J.E. and Rodrigues W.A. Jr. What is superluminal wave motion? *Sci. & Tech. Mag.*, August 1999, v.2 (accessed online from http://www.cptec.br/stm).
2. Moskowitz C. Strange particles may travel faster than light. Breaking Laws of Physics. In: LiveScience.com, 22 September 2011, http://news.yahoo.com/strange-particles-may-travel-faster-light-breaking-laws-192010201.html
3. Smarandache F. There is no speed barrier in the Universe. *Bulletin of Pure and Applied Sciences*, 1998, v. 17D (Physics), no. 1, 61 (accessed online from http://www.gallup.unm.edu/~smarandache/NoSpLim.htm).
4. Smarandache F. A hypothesis: There is no speed barrier in the Universe. Imprint: 01 Feb 1972, CERN Document Center, EXT-2004-133.
5. Einstein A., Podolsky B., and Rosen N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 1935, v.47, 777–780.
6. Bohm D. The Paradox of Einstein, Rosen, and Podolsky. *Quantum Theory*, 1951, 611–623.

7. Bell J.S. On the Einstein-Podolsky-Rosen Paradox. *Physics*, 1964, v.1, 195–200.
8. Weisstein E.W. Smarandache Hypothesis. In: *The Encyclopedia of Physics*, Wolfram Research, http://scienceworld.wolfram.com/physics/SmarandacheHypothesis.html

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), Zentralblatt MATH (Germany), 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, 200 College Road, Gallup, NM 87301, USA

Printed in the United States of America