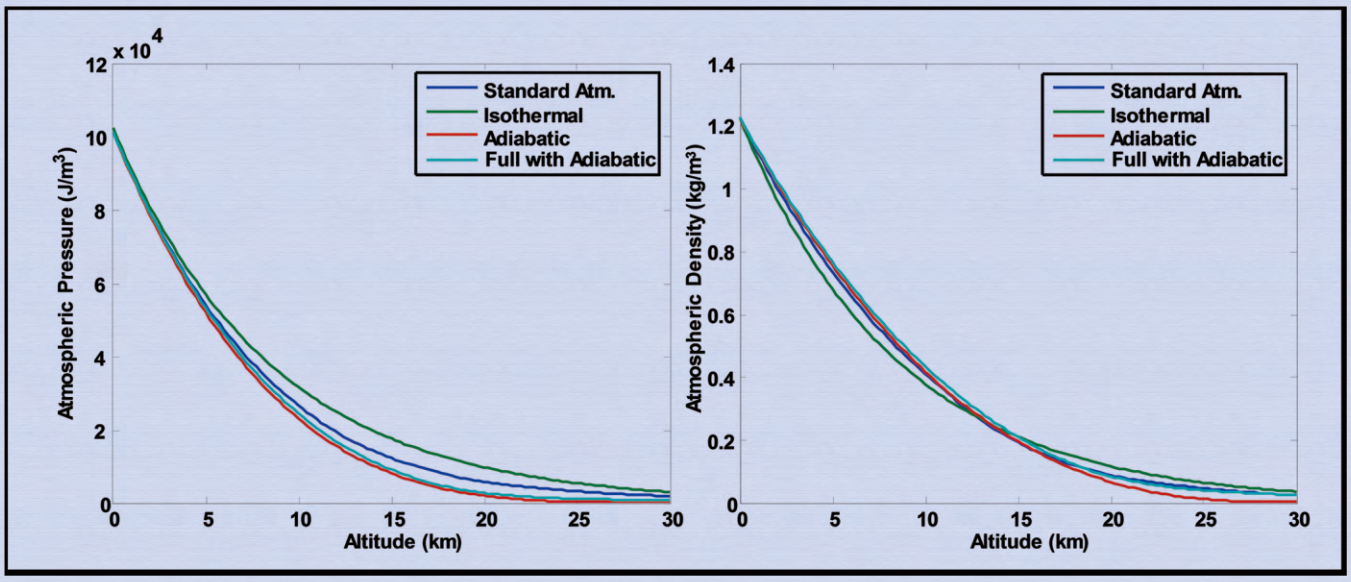
ISSN: 2153-1196 Volume 11, Number 5, May 2020

Scientific Research Publishing

Journal of Modern Physics



ISSN: 2153-1196

https://www.scirp.org/journal/jmp

Journal Editorial Board

ISSN: 2153-1196 (Print) ISSN: 2153-120X (Online)
https://www.scirp.org/journal/jmp

Editor-in-Chief

Editorial Board

Scientific Research Publishing

Journal of Modern Physics, 2020, 11, 593-765
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Table of Contents

Volume 11 Number 5 May 2020

The figure on the front cover is from the article published in Journal of Modern Physics, 2020, Vol. 11, No. 5, pp. 648-667 by Richard B. Holmes.

Journal of Modern Physics (JMP)
Journal Information

SUBSCRIPTIONS

The *Journal of Modern Physics* (Online at Scientific Research Publishing, https://www.scirp.org/) is published monthly by Scientific Research Publishing, Inc., USA.

Subscription rates:
Print: $89 per issue.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: jmp@scirp.org

Journal of Modern Physics, 2020, 11, 593-607
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

How to See Invisible Universes

Alexander Alexandrovich Antonov

Independent Researcher, Kiev, Ukraine
Email: telan@bk.ru

How to cite this paper: Antonov, A.A. (2020) How to See Invisible Universes. *Journal of Modern Physics*, **11**, 593-607.
https://doi.org/10.4236/jmp.2020.115039

Received: March 14, 2020
Accepted: April 27, 2020
Published: April 30, 2020

Abstract

The article shows that the special theory of relativity (STR) was not actually created in the 20th century, since: 1) the relativistic formulas presented therein are incorrect; 2) the relativistic formulas presented therein are explained incorrectly using the incorrect principle of light speed non-exceedance refuted in the article; 3) the relativistic formulas presented therein rise to incorrect conclusions about physical unreality of imaginary numbers and existence of only our visible universe. Moreover, the STR could not even have been created in the 20th century since: 1) experimental data on the six-dimensional space of our hidden Multiverse, which allowed to derive the correct relativistic formulas, were obtained by the WMAP and Planck spacecraft only in the 21st century; 2) the principle of physical reality of imaginary numbers, which allowed to refute the principle of not exceeding the speed of light and correctly explain the new relativistic formulas, was experimentally proved only in the 21st century. Therefore, the new relativistic formulas obtained given these circumstances allow to state that, in addition to our visible universe, there are numerous other mutually invisible universes, which together form the hidden Multiverse. Existence of invisible universes explains the phenomenon of dark matter and dark energy. Their existence also explains why antimatter in the hidden Multiverse does not annihilate with matter, and tachyons do not violate the principle of causality. The existence of these invisible universes can be proved by astronomical observations in the portals, since in them the constellations of the starry sky will differ from the constellations observed from existing observatories on Earth.

Keywords

Imaginary Numbers, Special Theory of Relativity, Dark Matter, Dark Energy, Dark Space, Multiverse, Hyperverse, Invisible Universes, Portals

1. Introduction

To see invisible universes means to ascertain their real physical existence and, therefore, discover them. Moreover, this means their indisputable discovery, since the universes can be surely identified by the constellations visible in their sky, like people can be identified by fingerprints. That is, constellations are a kind of passport of any universe that cannot be forged.

To see invisible universes, one needs to know how to do this, as they have not yet been discovered. However, in the existing version of the special theory of relativity (STR) [1] [2] [3] it is stated that in accordance with the principle of not exceeding the speed of light, other universes, except our visible universe, do not exist in nature. And this statement seems to have been even confirmed experimentally, as astronomers have never seen any other constellations in the sky, except those visible always. It is even considered that all the hypotheses of the Multiverse, proposed by scientists, some of which are given in [4]-[10], are unverifiable, *i.e.* they will neither now, nor in the distant future be in any way discovered by the inhabitants of the Earth.

So why then did the question referred to in the article headline arise? This is because ***other universes actually exist and can be seen by people*** [11].

2. Proof of Physical Reality of Concrete Imaginary Numbers

However, attempts to see invisible universes would make sense only provided that there is confidence in their existence and possibility of detection; that is, provided that it will be understood how this could be done. Then it would be proved that the STR denying this possibility and thereby the existence of Multiverses is wrong.

In order to make sure that the STR is wrong, we should recall that the principle of light speed non-exceedance implies not only denial of existence of Multiverses, but also denial of physical reality of imaginary numbers discovered five hundred years ago. And this turned out to be the whole point, since the authors of the STR did not know how to explain the relativistic formulas [12], in which all physical quantities, such as mass, time, etc. became imaginary at superluminal speeds. Neither the authors of the STR nor anyone before them knew what is, for example, $5i$ kilograms, $3i$ seconds, $2i$ meters, where $i=\sqrt{-1}$. However, physicists and mathematicians who lived before creation of the STR did not assert that imaginary numbers were not physically real. They admitted that they had no answer to this question. In contrast, the authors of the STR didn't get to admit this, since otherwise the STR would not have received recognition. That's why the STR turned out to require the postulate (*i.e.* the unproven assumption) on light speed non-exceedance, as it allowed the physical nature of concrete imaginary numbers not to explain.

However, Nobel Prize winner Stephen Weinberg was pretty clearly about such theories based on postulates: "*Scientific theories cannot be deduced by purely mathematical reasoning*". Therefore, the principle of light speed non-exceedance

in STR always caused doubt[1] in physicists. And in the 21st century physicists decided to verify experimentally the principle of light speed non-exceedance, as well as the assertion of physical unreality of imaginary numbers that followed from it. In 2011, the OPERA collaboration published a sensational report [13] stating that it had managed to refute the principle of light speed non-exceedance experimentally. But six months later the ICARUS collaboration refuted the OPERA experiment [14]. And thus impression was created that the principle of light speed non-exceedance is irrefutable.

However, it is possible that the creation of such an illusion was the true goal of the OPERA and ICARUS experiments. After all, back in 2008-2010 reports [15]-[20] were published on experimental studies of special processes in linear electric circuits, in which the general scientific principle of the physical reality of imaginary numbers was successfully proved. Therefore, the OPERA experiment has become unnecessary. Reports on further studies of the same purpose were published in [21]-[28]. But while the unsuccessful and unnecessary very expensive OPERA experiment was intensively touted, the physical community did not pay attention to simple alternative successful experimental evidence of the physical reality of imaginary numbers. Therefore, the assumption that the purpose of the OPERA and ICARUS experiments was not the search for scientific truth, but some other non scientific interests, does not seem to be unfounded.

Moreover, in publications [15]-[28] even three experimental proofs of the principle of physical reality of imaginary numbers were given:

- using transient oscillatory processes [19] [20] [24] [28], that allowed us to conclude that tsunamis would not have existed, church bells wouldn't have sounded and children's swing pushed by parents wouldn't have swung, if the STR statement about physical unreality of imaginary numbers were true;
- using resonant oscillatory processes [15] [16] [17] [18] [20] [21] [24] [28], that along with television, radiolocation and telecommunication could not have existed, if the STR statement about physical unreality of imaginary numbers were true;
- using Ohm's law in the interpretation of Steinmets [22] [23] [25] [26] [27] [28], proposed by him in 1893, which made it possible to refute the existing version of STR even before its creation.

And since the experiments presented in these publications, unlike the OPERA and ICARUS experiments, are simple and can be repeated in any radio engineering laboratory, ***the principle of physical reality of imaginary numbers is***

[1]The STR has been criticized by Oliver Heaviside, Nikola Tesla, Nobel Prize winner Albert Abraham Michelson, Nobel Prize winner Friedrich Wilhelm Ostwald, Nobel Prize winner Joseph John Thomson, Nobel Prize winner Svante August Arrhenius, Nobel Prize winner Philipp Eduard Anton von Lenard, Nobel Prize winner Alvar Gullstrand, Nobel Prize winner Wilhelm Carl Werner Otto Fritz Franz Wien, Nobel Prize winner Walther Hermann Nernst, Nobel Prize winner Ernest Rutherford, 1st Baron Rutherford of Nelson, Nobel Prize winner Johannes Stark, Nobel Prize winner Frederick Soddy, Nobel Prize winner Percy Williams Bridgman, Nobel Prize winner Edwin Mattison McMillan, Nobel Prize winner Hideki Yukawa, Nobel Prize winner Hannes Olof Gösta Alfven and many other outstanding scientists.

indisputably*[2] *proved by these experiments. The reluctance of the physical community to take into account the successful evidence of the physical reality of imaginary numbers set forth in these alternative radio engineering experiments speaks of the shortcomings in modern physical education.

Moreover, the denial in STO of the physical reality of the concrete imaginary numbers is equivalent to the denial of the physical reality corresponding to these numbers of the invisible and therefore still largely unknown world[3] which science will know about in the future. And this inhibits the development of science. Noting the great importance of imaginary numbers in the science, Sir Roger Penrose wrote: "*The system of complex numbers has a profound and timeless reality which goes beyond the mental constructions of any particular mathematician … They were put there neither by Cardano, nor by Bombelly, nor Wallis, nor Coates, nor Euler, nor Wessel, nor Gauss, despite the undoubted farsightedness of these, and other, great mathematicians; such set of magical properties was inherent in the very structure that they gradually uncovered …*"

Therefore, taking into account the principle of physical reality of imaginary numbers, although unacknowledged in physics, but experimentally proved in radio engineering, all physical theories and hypotheses should now be corrected.

3. Explanation of Physical Nature of Concrete Imaginary Numbers in the STR

Sometimes it is written that the STR has refuted the classical physics of Galileo, Newton and other great scientists of the past. But it is not so. In fact, as can be seen from Table 1 presenting the results of calculations using relativistic formulas, for example,

$$m = \frac{m_0}{\sqrt{1-(v/c)^2}} \tag{1}$$

$$\Delta t = \Delta t_0 \sqrt{1-(v/c)^2} \tag{2}$$

where m_0 is the rest mass of a physical body;

m is the relativistic mass of a moving physical body;

Δt_0 is the rest time of a physical body;

Δt is the relativistic time of a moving physical body;

v is the velocity of a physical body;

c is the speed of light,

Relativistic effects become noticeable only at speeds comparable to the speed of light. But such speeds are found only in space or at subatomic particle accelerators. There are no such speeds in everyday life, like in transport, in industry or in natural processes on Earth.

[2]What evidence could be more convincing than existence of, for example, TV-sets, tsunamis or children's swings that should not have existed, if the STR statement about physical unreality of imaginary numbers were true?

[3]However, an invisible world exists. For example, from the room in which we are now, the next room is not visible and, therefore, at that time it is an invisible world. Nevertheless, no one will deny its existence.

Table 1. Calculations, illustrating the absence of a relativistic effect in everyday human life in transport, in industry, in natural processes on Earth.

v/c	0.0100	0.1000	0.9000	0.9900	0.9990	0.9999
m/m_0	1.00005	1.00504	2.29416	7.08881	22.36627	70.71245
$\Delta t/\Delta t_0$	0.99995	0.99498	0.43589	0.14167	0.04471	0.01414

Consequently, the existing version of SRT is not only not quite true, but it is not needed by people in everyday life. And still need classical physics.

But in science, for example, in astrophysics, STR is needed. More precisely, we need a corrected version of it. The denial by the existing version of the SRT of the principle of physical reality of imaginary numbers, as has just been shown, is its mistake. Therefore, the error must be corrected, *i.e.* it must be taken into account that relativistic formulas (1), (2), etc. must be explainable both in the range of $0 \le v < c$ and in the range of $c \le v < \infty$. However, the graphs of formulas (1) and (2) in **Figure 1(a)**, **Figure 1(b)** shows that the formulas are incorrect, since their graphs in the ranges $0 \le v < c$ and $c \le v < \infty$ are of significantly different forms. Therefore, they are inexplicable.

Analysis of their possible explainable options has showed that the graphs of relativistic formulas should be written as shown in **Figure 1(c)**, **Figure 1(d)**. Therefore corrected relativistic formulas should take the following form

$$m = \frac{m_0 i^q}{\sqrt{1-\left(v/c-q\right)^2}} = \frac{m_0 i^q}{\sqrt{1-\left(w/c\right)^2}} \tag{3}$$

$$\Delta t = \Delta t_0 i^q \sqrt{1-\left(v/c-q\right)^2} = \Delta t_0 i^q \sqrt{1-\left(w/c\right)^2} \tag{4}$$

where $q = \lfloor v/c \rfloor$ is the "floor" function of argument v/c, whose values[4] correspond to different parallel[5] universes;

$w = v - qc$ is the local velocity for each universe, which can take values only in the range of $0 \le w < c$;

v is the velocity measured for our visible universe.

Thus, formulas (3) and (4) can already be interpreted in a comprehensible manner. Their q parameter can be considered as an extra spatial dimension[6]. Thus, the value $q = 0$ would correspond to our visible universe (as $i^0 = 1$); and the value $q = 1$ would correspond to another universe (as $i^1 = i$) invisible for us, given that $v > c$, as it is beyond the horizon of events. It shall, therefore, be referred to as a tachyon universe, just like subatomic particles moving at superluminal speeds. For the same reasons, our universe shall be referred to as tardyon. The value $q = 2$ would correspond to a tardyon antiverse (as $i^2 = -1$), the value $q = 3$ would correspond to a tachyon antiverse (as $i^3 = -i$), the value $q = 4$ would correspond to another[7] tardyon universe (as $i^4 = 1$), the value $q = 5$ would correspond to another tachyon universe (as $i^5 = i$ for it), etc.

[4]And it takes non-integral values in the portals considered below, in which, their value q changes for one from entry to exit under the influence of physical factors not yet studied.

[5]Called so because they do not intersect, despite their infinity.

[6]Like numbers of apartments in an apartment building.

[7]There are about twenty such invisible universes in the hidden Multiverse, as shown below.

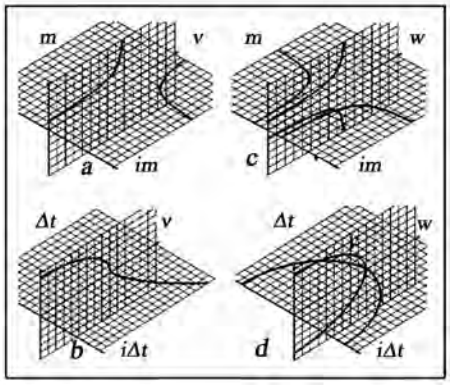


Figure 1. Graphs of functions (1), (2) corresponding to the existing wrong version of the STR, and (3), (4) corresponding to its corrected version.

Thus, ***concrete imaginary numbers in the corrected relativistic formulas (3), (4), etc. correspond to mutually invisible parallel universes, together forming the Multiverse, which we therefore call the hidden Multiverse*** [29] [30] [31].

However, the invisible universes mentioned are not a gift in science, since they are known even less than, for example, dark matter and dark energy. Moreover, despite the logically unimpeachable arguments presented in [32] [33] confirming real physical existence of invisible universes, there is still no[8] complete certainty about this due to the lack of indisputably conclusive experiments.

4. Explanation of Dark Matter and Dark Energy by Existence of Invisible Universes of the Hidden Multiverse

However, existence of dark matter and dark energy that are as much incomprehensible as invisible universes has already been experimentally proven: dark matter in 1932-33 by Jan Hendrick Oort and Fritz Zwicky, and dark energy in 1998-99 by Saul Perlmutter, Brian Schmidt and Adam Riess, who were awarded the Nobel Prize for this discovery. Emphasizing the outstanding importance of these discoveries, Nobel Prize winner Adam Riess wrote: "*Humanity is on the verge of a new physics of the Universe. Whether we want it or not, we will have to accept it.*" And about this new physics in the article is discussed below.

Dark matter and dark energy [34] [35] [36] [37] are called dark for their obscurity. It is unclear why they are absolutely invisible in all ranges of electromagnetic oscillations and can only be detected indirectly by their gravitational manifestations. It is unclear why neither molecules, nor atoms, nor subatomic particles are found in dark matter and dark energy, though their total mass/energy is more than 20 times greater than the mass/energy of our visible universe. Many other things are as well unclear. Therefore, their knowledge in modern physics is a task of paramount importance.

Professor Michio Kaku commented on this situation as follows: "*Of course, a whole bunch of Nobel Prizes is waiting for the scientists who can reveal the se-*

[8]However, it is written below how to get them.

crets of the dark energy and dark matter".

The phenomenon of dark matter and dark energy is currently investigated very intensively. But it is still far from being fully explained, since all the studies are conducted only in terms of the hypothesis of the visible Monoverse that follows from the erroneous version of the STR. And in this case we cannot but agree with the Nobel Prize laureate Albert Einstein, who argued: "*Insanity*: *doing the same thing over and over again and expecting different results*". In this regard, it is pertinent to recall Sir Isaac Newton's opinion: "*No great discovery was ever made without a bold guess*".

Therefore, assuming that there is a hidden Multiverse rather than a Monoverse, the phenomenon of dark matter and dark energy can be explained as follows [38]-[45]:

- dark matter and dark energy do not actually have any real physical contents; they are some kind of image (but not electromagnetic) analogous to an optical, something like a shadow evoked by existence of invisible parallel universes;
- herewith, the phenomenon of dark matter is evoked by invisible parallel universes adjacent to our visible universe, and the phenomenon of dark energy is evoked by other invisible parallel universes of the hidden Multiverse;
- therefore, no material contents, such as molecules, atoms or subatomic particles, will ever be found in dark matter and dark energy;
- dark matter and dark energy are invisible because other parallel universes of the hidden Multiverse, except our visible universe, are invisible.

Such approach allows getting an idea of the structure of the hidden Multiverse, based on experimental data from WMAP spacecraft that was launched by the National Aeronautics and Space Administration (NASA) in 2001 and operated until 2009, as well as data from the Planck spacecraft that was launched by the European Space Agency (ESA) in 2009 and operated until 2013. Thus, according to the WMAP spacecraft date [46], the entire universe (in fact, the entire hidden Multiverse) is 4.6% composed of baryonic matter, 22.4% of dark matter, and 73.0% of dark energy. And according to later data from the Planck spacecraft [47], the entire universe (in fact, the entire hidden Multiverse) is 4.9% composed of baryonic matter, 26.8% of dark matter, and 68.3% of dark energy. So, assuming that over billions of years of existence mass/energy of all invisible universes has substantially averaged due to exchange of their material contents through portals[9], it can be argued that:

- according to the WMAP spacecraft data the hidden Multiverse is formed by **100%/4.6% = 21.74** universes, and according to the Planck spacecraft data they are **100%/4.9% = 20.41** universes, *i.e.* real number of the universes is supposedly **20 … 22**;
- according to the WMAP spacecraft data **22.4%/4.6% = 4.87** universes correspond to the dark matter phenomenon, and according to the Planck spacecraft data they are **26.8%/4.9% = 5.47** universes, *i.e.* real number of the un-

[9]That will be discussed below.

iverses is supposedly **5 … 6**;

- according to the WMAP spacecraft data **73.00%/4.6% = 15.87** universes correspond to the dark energy phenomenon, and according to the Planck spacecraft data they are **68.3%/4.9% = 13.94** universes, *i.e.* real number of the universes is supposedly **14 … 16**.

And ***such a result in the 20th century was impossible to guess by any postulate***.

But it is not difficult to notice that the results obtained as a result of calculations do not correspond to the corrected relativistic formulas (3) and (4). Indeed, according to formulas (1) and (2). Our tardyon universe in the hidden Multiverse should have only two adjacent universes in one extra dimension, rather than five or six universes, as follows from the above calculations. Therefore, it follows from the WMAP and Planck spacecraft data, that the number of extra dimensions should be three.

Thus, formulas (3) and (4) must be corrected again as follows

$$\begin{aligned} m &= \frac{m_0 \left(i_1\right)^{q-q_0} \left(i_2\right)^{r-r_0} \left(i_3\right)^{s-s_0}}{\sqrt{1-\left[v/c-\left(q+r+s-q_0-r_0-s_0\right)\right]^2}} \\ &= \frac{m_0 \left(i_1\right)^{q-q_0} \left(i_2\right)^{r-r_0} \left(i_3\right)^{s-s_0}}{\sqrt{1-\left(w/c\right)^2}} \end{aligned} \tag{5}$$

$$\begin{aligned} \Delta t &= \Delta t_0 \left(i_1\right)^{q-q_0} \left(i_2\right)^{r-r_0} \left(i_3\right)^{s-s_0} \sqrt{1-\left[v/c-\left(q+r+s-q_0-r_0-s_0\right)\right]^2} \\ &= \Delta t_0 \left(i_1\right)^{q-q_0} \left(i_2\right)^{r-r_0} \left(i_3\right)^{s-s_0} \sqrt{1-\left(w/c\right)^2} \end{aligned} \tag{6}$$

where q_0, r_0, s_0 are the coordinates of our visible universe in the hidden Multi-verse:

v is the velocity measured from our tardyon universe;

c is the speed of light;

$w = v - \left(q+r+s-q_0-r_0-s_0\right)c$ is the local velocity for the universe corresponding to the coordinates q, r, s of the universe, which can take values only in the range of $0 \le w \le c$.

Herewith, three imaginary units i_1, i_2, i_3 that correspond to three extra dimensions and are present in hypercomplex numbers called quaternions [48], are interconnected by the relations

$$i_1^2 = i_2^2 = i_3^2 = 1 \tag{7}$$

$$i_1 i_2 i_3 = i_2 i_3 i_1 = i_3 i_1 i_2 = -1 \tag{8}$$

$$i_1 i_3 i_2 = i_2 i_1 i_3 = i_3 i_2 i_1 = 1 \tag{9}$$

Consequently, the ***structure of the six-dimensional space*** (***see*** Figure 2) ***of the hidden Multiverse can be described by the formula*** $f_{q,r,s}\left(x,y,z\right) + i_1 q + i_2 r + i_3 s$ ***in which the summand*** $i_1 q + i_2 r + i_3 s$ ***determines the coordinates of the corresponding invisible universe, and the summand*** $f_{q,r,s}\left(x,y,z\right)$ ***determines the distribution in the coordinates*** x, y, z ***of its material contents in this universe.***

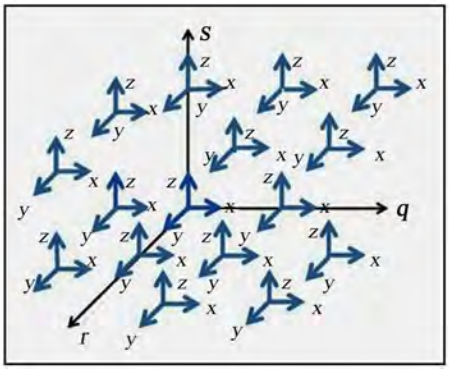


Figure 2. Six-dimensional space of the hidden Multiverse.

Professor Lisa Randall wrote about a similar situation: "*We probably live in a three-dimensional gap of higher dimensional space*". And her assumption, as seen, was confirmed.

Figure 3 presents an example of such a quaternion structure of the hidden Multiverse corresponding to the results of mathematical processing of WMAP and Planck data and containing twenty-one parallel universes. As can be seen, tardyon universes and antiverses alternate with tachyon universes and antiverses in such an open screw structure. Pursuant to the formula (7), they are interconnected by numerous bidirectional portals denoted by bidirectional arrows. And tachyon universes and antiverses are interconnected by numerous unidirectional portals denoted by unidirectional arrows pursuant to the formulas (8) and (9). Herewith, invisible tachyon universes and antiverses adjacent to our tardyon universe evoke the phenomenon of dark matter. And the remaining invisible universes of the hidden Multiverse evoke the phenomenon of dark energy.

Thus, two obscure concepts "invisible universes" and "dark matter and dark energy" let us understand their meaning in joint consideration. Dark matter and dark energy turn out to be the result of recording gravitational manifestations of invisible parallel universes of the hidden Multiverse in our visible universe.

Advantage of such hypothesis of dark matter and dark energy is that, in addition to explanation of dark matter and dark energy inexplicable until recently, it also explains the inexplicable issues regarding the existence of antimatter and tachyons [49]. As can be seen from **Figure 3**, antimatter exists in antiverses and it does not annihilate with matter because tardyon universes and antiverses alternate with tachyon universes and antiverses. Moreover, there are even several tardyon and tachyon matters and antimatters in the hidden Multiverse. And tachyons that do not violate the principle of causality exist in tachyon rather than in tardyon universes and antiverses.

5. Comments on the Existing Version of the STR

The results obtained are so important that they need additional comments. STR actually consists of two large sections, in the first of which relativistic formulas are derived, and in the second these formulas are explained.

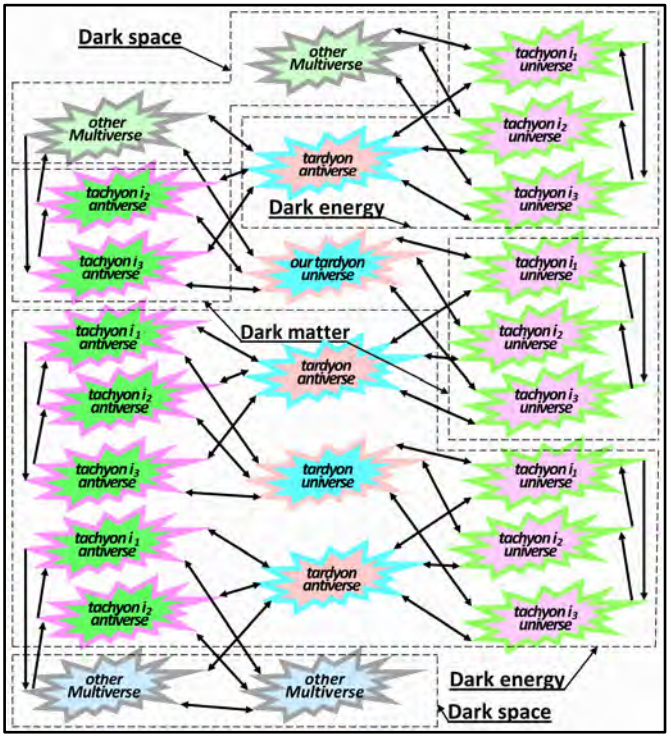


Figure 3. Example of the quaternion structure of the hidden Multiverse.

And it turned out that in both sections of the existing version of the STR, incorrect results were obtained:

- in the first section, instead of the valid relativistic formulas (5) and (6), the incorrect formulas (1) and (2) are obtained;
- in the second section, from the wrong relativistic formulas (1) and (2), erroneous conclusions were drawn about the physical unreality of imaginary numbers and the existence in nature of our only visible universe.

Thus, since the existing version of the STR is incorrect, we have to conclude that ***in the 20th century STR was not created*** [50] [51]. ***Moreover, it could not be created at that time, since Albert Einstein was ahead of his time*** the experimental data necessary for the derivation of relativistic formulas (5), (6) and other spacecraft WMAP and Planck were obtained only in the 21st century. And the principle of the physical reality of imaginary numbers, necessary for the interpretation of these formulas, was also proved only in the 21st century.

Nevertheless, the very statement of the task of creating STR is an outstanding scientific achievement of Albert Einstein. Without an intermediate result in the form of formulas (1), (2), the final result in the form of formulas (5), (6) would not have been obtained. And the basic principles for solving this problem are described in this article.

6. Explanation of Dark Space by Existence of Invisible Universes of the Hyperverse

As can be seen from **Figure 3**, our visible universe and some invisible universes

in the hidden Multiverse are connected not only with each other but also with invisible universes of other Multiverses through unidirectional and bidirectional portals denoted by single-sided and double-sided arrows [52] [53] [54]. These other Multiverses relevant to the phenomenon of dark space [55] [56] together with our hidden Multiverse form the Hyperverse (see **Figure 4**).

And if this were not so, then the hidden Multiverse would have to have edges that contradict common sense and cannot be explained in any way, as, for example, archaic conception about shape of Earth.

But the phenomenon of dark space with WMAP and Planck spacecraft is not registered. However, it can be registered with the studies described below of invisible universes in portals.

7. Verifiability of Invisible Universes

Now it's time to answer the question formulated in the article headline "How to see invisible universes?" The question can also be formulated in another way, since it is not completely clear what means to see a universe and why it is necessarily.

This is not necessarily, but such evidence of existence of invisible universes is the most understandable and convincing. Invisible universes can be compared with rooms of our house that are invisible to us from the room we are in now. It is clear that to see invisible rooms one has to come to the doors and look into or enter the rooms. In the same way, ***to see invisible universes, one has to penetrate the portals and look at the starry sky. It must be different in other universes. Other constellations must be observed on it*** [11] [55] [56]. And these differences would apparently be the greater the further one enters the portals by foot or by vehicle[10]. The starry sky would be completely different after leaving portals and getting into other universes, which would become visible to us, unlike our universe.

Besides, it is clear that satisfaction of our curiosity in this way would be complicated by two things. The first is that the portals are invisible, and the second is that we haven't been invited to other universes through these portals. However it is possible that some portals, which are very numerous on the Earth [57], are not protected from our visit. Therefore, one can probably try to visit them.

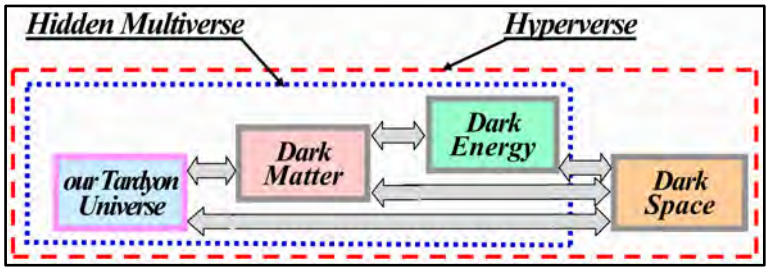


Figure 4. Structure of the Hyperverse.

[10]Therefore, there is no need to fly in portals and through portals by rockets.

In order not to get lost in the portals and to have a guaranteed opportunity to return to Earth located in our universe after visiting them, it is desirable to have some navigation devices, a kind of a marine compass. These devices of portal navigation should indicate the direction of movement from portal entries to exits and vice versa. Creation of such devices is possible, because with distance from portal entry radio signal strength, e.g, for mobile communication, would decrease and after leaving the portal for the adjacent universe, the signal will completely disappear.

However, it is not necessary to penetrate the portals far off their entries and use navigation devices make certain of existence of invisible universes. The fact is that as you move through the portal from our visible universe to one of the neighboring invisible universes, the map of the earth's starry sky will smoothly transform into a map of the starry sky of another universe. At that the maps of the starry sky in our and in any of the neighboring universes, of course, differ in an extreme way. ***Therefore, even with shallow penetration into the portals, it will be found that the constellations above the head of the researcher will be so markedly different from the constellations observed in Earth observatories that even amateur astronomers can probably detect and register these changes with their telescopes.*** And of course, these differences will be noticeably larger than the differences recorded in 1919 in a similar experiment by Sir Arthur Stanley Eddington [58].

8. Conclusions

The article shows that the existing version of SRT is incomplete and, therefore, in the 20th century this theory has not yet been created. Moreover, SRT in the 20th century could not have been created, since the data obtained by the WMAP and Planck spacecraft only in the 21st century were necessary to derive the correct relativistic formulas. And for their correct interpretation, the principle of the physical reality of imaginary numbers was needed, experimentally proved in the 21st century. The article also proposes and explains the corrected relativistic formulas of the STR, from which it follows that there is a hidden Multiverse formed by invisible parallel universes.

It is explained how the existence of invisible universes can be convincingly proved experimentally. For this, it is enough to use the obvious truth, the starry sky of other universes contains other constellations. Therefore, in portals, which are transitional zones between neighboring universes, the constellations of the starry sky should also be different. And this starry sky in the portals, which on Earth are, at least, some of the many so-called anomalous zones, can be seen and registered. In the entire history of the existence of astronomy, it has not had a task more interesting and more important than the discovery and study of invisible universes.

Thus, the hypothesis of the hidden Multiverse and Hyperverse set forth in the article is verifiable, and the invisible universes are real and can be seen.

Acknowledgements

The author is grateful for participation in the discussion of the paper to Olga Ilyinichna Antonova, whose criticism and valuable comments contributed to improvement of the article.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Einstein, A. (1920) Relativity: The Special and General Theory. H. Holt and Company, New York.

[2] Bohm, D. (2006) The Special Theory of Relativity. Routledge, Abingdon-on-Thames.

[3] Hawking, S.W. and Penrose, R. (2010) The Nature of Space and Time. Princeton University Press, Princeton. https://doi.org/10.1515/9781400834747

[4] Kaku, M. (2006) Parallel Worlds: A Journey through Creation, Higher Dimensions, and the Future of the Cosmos. Anchor, New York.

[5] Vilenkin, A. (2007) Many Worlds in One: The Search for Other Universes. Hill and Wong, New York.

[6] Carr, B. (2009) Universe or Multiverse? Cambridge University Press, Cambridge.

[7] Gribbin, J. (2010) In Search of the Multiverse: Parallel Worlds, Hidden Dimensions, and the Ultimate Quest for the Frontiers of Reality. Wiley & Sons Inc, Hoboken.

[8] Greene, B. (2011) The Hidden Reality: Parallel Universes and the Deep Laws of the Cosmos. Vintage, New York.

[9] Hawking, S. and Mlodinow, L. (2012) The Grand Design. Bantam, New York.

[10] Tegmark, M. (2014) Our Mathematical Universe: My Quest for the Ultimate Nature of Reality. Vintage, New York.

[11] Antonov, A.A. (2019) *Journal of Modern Physics*, **10**, 1006-1028. https://doi.org/10.4236/jmp.2019.108067

[12] Antonov, A.A. (2014) *American Journal of Scientific and Industrial Research*, **5**, 40-52.

[13] Adam, T., Agafonova, N., Aleksandrov, A., *et al.* (2012) *Journal of High Energy Physics*, **2012**, Article No. 93. https://doi.org/10.1007/JHEP10(2012)093

[14] Antonello, M., Baibussinov, B., Boffelli, F., *et al.* (2012) Precision Measurement of the Neutrino Velocity with the ICARUS Detector in the CNGS Beam.

[15] Antonov, A.A. (2008) *European Journal of Scientific Research*, **21**, 627-641.

[16] Antonov, A.A. (2009) *European Journal of Scientific Research*, **28**, 193-204.

[17] Antonov, A.A. (2010) *American Journal of Scientific and Industrial Research*, **1**, 342-349. https://doi.org/10.5251/ajsir.2010.1.2.342.349

[18] Antonov, A.A. (2010) *International Journal of Pure and Applied Sciences and Technology*, **1**, 1-12.

[19] Antonov, A.A. (2010) *General Mathematics Notes*, **1**, 11-16.

[20] Antonov, A.A. (2013) Unpredictable Discoveries. Lambert Academic Publishing, Saarbrücken.

[21] Antonov, A.A. (2015) *General Mathematics Notes*, **31**, 34-53.

[22] Antono, A.A. (2015) *American Journal of Electrical and Electronics Engineering*, **3**, 124-129.

[23] Antonov, A.A. (2015) *Global Journal of Physics*, **2**, 145-149.

[24] Antonov, A.A. (2016) *General Mathematics Notes*, **35**, 40-63.

[25] Antonov, A.A. (2016) *PONTE*, **72**, 131-142. https://doi.org/10.21506/j.ponte.2016.7.9

[26] Antonov, A.A. (2016) *Journal of Modern Physics*, **7**, 2299-2313. https://doi.org/10.4236/jmp.2016.716198

[27] Antonov, A.A. (2016) *International Review of Physics*, **10**, 31-35. https://www.praiseworthyprize.org/jsm/index.php?journal=irephy&page=article&op=view&path%5B%5D=18615

[28] Antonov, A.A. (2017) *Norwegian Journal of Development of the International Science*, **6**, 50-63.

[29] Antonov, A.A. (2015) *International Journal of Advanced Research in Physical Science*, **2**, 25-32.

[30] Antonov, A.A. (2016) *Journal of Modern Physics*, **7**, 1933-1943. https://doi.org/10.4236/jmp.2016.714170

[31] Antonov, A.A. (2017) *Natural Science*, **9**, 43-62. https://doi.org/10.4236/ns.2017.93005

[32] Antonov, A.A. (2016) *Global Journal of Science Frontier Research*: *A Physics and Space Science*, **16**, 4-12.

[33] Antonov, A.A. (2017) *Global Journal of Science Frontier Research*. http://blog.gjsfr.org/2017/02/verifiable-hidden-multiverse.html

[34] Ruiz-Lapuente, P. (2010) Dark Energy: Observational and Theoretical Approaches. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139193627

[35] Amendola, L. and Tsujikawa, S. (2010) Dark Energy: Theory and Observations. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511750823

[36] Bertone, G. (2013) Particle Dark Matter: Observations, Models and Searches. Cambridge University Press, Cambridge.

[37] Sanders, R.H. (2014) The Dark Matter Problem: A Historical Perspective. Cambridge University Press, Cambridge.

[38] Antonov, A.A. (2015) *American Journal of Modern Physics*, **4**, 1-9. https://doi.org/10.11648/j.ajmp.20150401.11

[39] Antonov, A.A. (2015) *International Journal of Physics*, **3**, 84-87.

[40] Antonov, A.A. (2015) *Cosmology*, **19**, 40-61.

[41] Antonov, A.A. (2015) *American Journal of Modern Physics*, **4**, 180-188. https://doi.org/10.11648/j.ajmp.20150404.14

[42] Antonov, A.A. (2015) *Optics*, **4**, 43-47.

[43] Antonov, A.A. (2016) *Frontiers of Astronomy, Astrophysics and Cosmology*, **2**, 1-9.

[44] Antonov, A.A. (2016) *Journal of Modern Physics*, **7**, 1228-1246. https://doi.org/10.4236/jmp.2016.710111

[45] Antonov, A.A. (2017) *Journal of Modern Physics*, **8**, 567-582. https://doi.org/10.4236/jmp.2017.84038

[46] Hinshaw, G., Larson, D., Komatsu, E., *et al.* (2013) Nine Year Wilkinson Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results.

[47] Adam, R., Ade, P.A.R., Aghanim, N., *et al.* (2015) Plank 2015 Results. 1. Overviev of Products and Scientific Results.

[48] Kantor, I.L. and Solodovnikov, A.S. (1989) Hypercomplex Numbers. Springer, Verlag, Berlin. https://doi.org/10.1007/978-1-4612-3650-4

[49] Antonov, A.A. (2016) *PONTE*, **72**, 288-300. https://doi.org/10.21506/j.ponte.2016.9.22

[50] Antonov, A.A. (2019) *Journal of Russian Physical-Chemical Society*, **91**, 57-94. (In Russian) http://rusphysics.ru/files/Antonov.91-1.pdf

[51] Antonov, A.A. (2020) *Journal of Modern Physics*, **11**, 324-342. https://doi.org/10.4236/jmp.2020.112020

[52] Antonov, A.A. (2015) Where to Look for Alien Civilisations. Cosmology. Commentaries: Stephen Hawking's Aliens. The Search for Intelligent Extraterrestrial Life. Project Breakthrough Liste. http://cosmology.com/Aliens1.html

[53] Antonov, A.A. (2016) How Portals of the Invisible Multiverse Operate. SciencePG Frontiers. Research Blog. http://www.sciencepublishinggroup.com/news/sciencepgfrontiersinfo?articleid=147

[54] Antonov, A.A. (2016) *Philosophy and Cosmology*, **6**, 11-27. (In Russian)

[55] Antonov, A.A. (2018) *Journal of Modern Physics*, **9**, 14-34. https://doi.org/10.4236/jmp.2018.91002

[56] Antonov, A.A. (2018) *Natural Science*, **10**, 11-30. https://doi.org/10.4236/ns.2018.101002

[57] Chernobrov, V.A. (2000) Encyclopedia of Mysterious Places of the Earth. Veche Publishing, Bucharest. (In Russian)

[58] Dyson, F.W., Eddington, A.S. and Davidson, C. (1920) *Philosophical Transactions of the Royal Society A*: *Mathematical, Physical and Engineering Sciences*, **220**, 291-333. https://doi.org/10.1098/rsta.1920.0009

Journal of Modern Physics, 2020, 11, 608-615
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

The Beta-Decay Induced by Neutrino Flux

Boris V. Vasiliev

Dubna, Russia
Email: bv.vasiliev@yandex.com

How to cite this paper: Vasiliev, B.V. (2020) The Beta-Decay Induced by Neutrino Flux. *Journal of Modern Physics*, **11**, 608-615.
https://doi.org/10.4236/jmp.2020.115040

Received: April 5, 2020
Accepted: May 5, 2020
Published: May 8, 2020

Abstract

Previously, it has been repeatedly suggested that the radioactive decay of nuclei is not a random phenomenon and can occur under the influence of a neutrino flux. Our recent experiment [1] showed that the neutrino flux created by a nuclear reactor affects the decay of nuclei in an isolated source $^{90}Sr/^{90}Y$, whose beta-electrons have an average energy of the order of 1 MeV. This paper presents the results of searching for the effect of the neutrino flux generated by a pulsed nuclear reactor on the rate of decay of ^{63}Ni nuclei. These nuclei were chosen as the object of research due to the fact that they have a low energy of beta-electrons of the order of 50 keV. Measurements have shown that the same flux of reactor neutrinos has approximately an order of magnitude stronger effect on ^{63}Ni nuclei than on the previously studied nuclei $^{90}Sr/^{90}Y$.

Keywords

Nuclear Reactor, Neutrino Flux, Beta-Decay, Isotope ^{63}Ni

1. Introduction

The discovery of radioactivity raised the question: is radioactive decay a random phenomenon?

The new physics of the early twentieth century gave the phenomenon of radioactive decay a quantum-mechanical explanation. N. Bohr and other creators of quantum mechanics explained it by the laws of tunnel phenomena and considered its absolute randomness to be proved.

The power of the quantum mechanics apparatus and the beauty of the new approach have swayed the physical community to the side of the anti-determinists. Currently, this view is dominant and radioactive decay is considered a purely random phenomenon.

But some other physicists from the very beginning of discussion did not agree with this approach.

A. Einstein expressed his categorical rejection of the new point of view with his famous metaphor: "God doesn't play dice!"

From a deterministic point of view, radioactive decay must have been caused by some hitherto unknown external causes.

The neutrino flux is very well suited for the description of the external causes that can induce beta-decay.

The hypothesis that the impact of the neutrino flux may be the cause of beta-decay of radioactive nuclei has been repeatedly expressed earlier by various researchers [2] [3] [4].

It is important that the phenomenon of beta-decay of neutron is energetically advantageous, since the mass of neutron is greater than the total mass of proton and electron.

All main parameters of neutron, its mass, spin, magnetic moment, decay energy, can be calculated with satisfactory accuracy in the framework of the electromagnetic model [5] [6] [7]. This model predicts the characteristic properties of neutron excited states [6] that it should possess. In high-energy physics, particles with these properties are commonly called hyperons.

(Only the neutron lifetime is not computable).

Electromagnetic forces that, according to the electromagnetic model, bind proton and electron to form neutron do not carry a mechanism for degradation of the combined state.

In this model, there is no internal cause that could induce the lose stability of neutron and its decay.

Therefore, in order to cause the decay of neutrons or β-active nuclei, they must be affected from the outside by the neutrino flux.

In previous paper, it was shown that the flux of reactor neutrinos does have a small effect on the rate of beta-decay in the source ^{90}Sr/^{90}Y [1].

In this paper, it is shown that the same flux of reactor neutrinos has approximately an order of magnitude stronger effect on the decay of ^{63}Ni nuclei, characterized by low energy beta-electrons.

1.1. The Neutron Lifetime

Multiple studies of the neutron lifetime have shown one feature. Neutrons in the last century at very accurate measurements made on relatively small reactors lived noticeably longer than neutrons lived near the powerful Grenoble reactor at the beginning of the XXI century.

The difference in life time of these neutrons in these changes significantly went beyond the limits of experimental errors [8]. This result can be explained by the fact that the powerful neutrino flux of the Grenoble reactor causes accelerated neutron decay.

1.2. The Effect of Solar Neutrinos on β-Decay of Nuclei

A clear result was obtained by E. Fishbach and J. Jenkins [3].

They showed that according to the data of experiments conducted in various laboratories around the world, the rate of nuclei decay depends on the distance between the Sun and the Earth. The authors explained this effect by the influence of the solar neutrino flux.

Accordingly to measurement data [9], the sunny neutrino flux is equal approximately

$$\Phi_{\odot} \approx 6\times 10^{10}\, \nu/\mathrm{cm}^2 \cdot \mathrm{s} \tag{1}$$

To explain the measured modulation depth, it can be assumed that the solar neutrino flux is approximately only 1/30 of the total neutrino flux that affects the decay of nuclei in our laboratories, and this total neutrino flux from all cosmic sources is equal to

$$\Phi_{cosm} \approx 30 \cdot \Phi_{\odot} \approx 2\times 10^{12}\, \nu/\mathrm{cm}^2 \cdot \mathrm{s} \tag{2}$$

2. The Effect of Reactor Neutrinos on β-Decay

Experiments on a Stationary Reactor

In order to detect the effect of neutrino flux on the beta-decay phenomenon, a group of researchers led by E. Fishbach conducted experiments on the High Flux Isotope Reactor of Oak Ridge National Laboratory [10].

This reactor had a high power

$$W_{HFIR} \approx 85\ \mathrm{MW} \approx 1.4\times 10^{21}\ \mathrm{MeV/s}. \tag{3}$$

Given that the fission of a single nucleus of reactor fuel releases approximately 200 MeV of energy [11], this reactor produced approximately 10^{19} of nuclear fission per second.

If we assume that one act of fission produces 6 neutrino [11], it turns out that at a conditional distance of 10 meters from the core, this reactor created a neutrino flux

$$\Phi_{HFIR} \approx 3\times 10^{12}\, \nu/\mathrm{cm}^2 \cdot \mathrm{sec} \tag{4}$$

Thus, the neutrino flux generated by this reactor was approximately equal to the flux of cosmic neutrino (Eq. (2)).

Therefore, according to the authors, switching the reactor on or off should have greatly changed the rate of decay in the beta sources under study.

However, despite a very serious approach to the processing of experimental results, the authors did not find any effect of such a powerful reactor on the beta-decay of a number of studied nuclei [10].

How this discrepancy in research results ([3] and [10]) can be explained? The possible reason for this is due to the difference in the spectra of cosmic and reactor neutrinos.

Most reactor neutrinos generated by fission fragments have an energy of the order of 1 MeV or less. At the same time among the cosmic neutrinos, there are particles with very high energy.

In this regard, it can be assumed that the average flow of reactor neutrinos

may have a smaller effect on beta-decay than the same-valued flow of cosmic neutrinos.

In this regard, in order to notice the effects of reactor neutrinos on beta-decay, it is necessary to take a different approach to this experiment.

3. Experiments on a Pulsed Reactor

3.1. The Pulse Reactor Efficiency

For this experiment, we used the IBR-2 reactor (Dubna) as a neutrino source. It has an average power of 1.6 MW and operates in pulse mode. The interval between the flashes was about 200 ms.

The duration of the flash itself is approximately 0.3 ms.

The estimate shows that the neutrino flux generated by this reactor during the flash of its intensity is approximately equal to the flow of a powerful stationary reactor or the flow of cosmic neutrinos.

Using a pulse reactor makes it possible to accumulate information synchronously with the flash. This is an important advantage because over a long time of measurement, it is possible to sum up the data obtained in the on-off mode of the reactor.

This method of data accumulation is equivalent to the increasing of the reactor power by several orders of magnitude.

3.2. The Reactor Neutrinos Generation

During the fission reaction of nuclear fuel, β-active nuclei are born in the form of fission fragments. Their decay leads to the generation of neutrinos.

However, almost all fission fragments of nuclear fuel have a half-life longer than the time period separating the flares of our reactor. Therefore, in our case, they can be considered long-lived. Their decays are not "tied" to a specific flash of the reactor and can only give a slight increase in the background in the accumulation mode.

There are very few short-lived isotopes with a half-life of $T_{1/2} < 200$ ms in nature. One of these isotopes, ^{12}B, has a half-life of about 20 ms [12].

The probability of its formation as a plutonium fission fragment is unknown and should probably be very small.

But it is a great success that the neutrinos formed during its decay have a high boundary energy-up to 13.3 MeV [12].

It is likely that the beta-decay of ^{12}B creates those neutrinos whose effect on the beta-decay at the source of ^{90}Sr/^{90}Y was registered earlier (**Figure 1**) [1].

The measurement time is 3 days.

4. Experimental Setup and Measurement Results

A new series of experiments was carried out on the same installation that had previously been used for measurements effects of the neutrino flux on the decay of nuclei ^{90}Sr/^{90}Y [1].

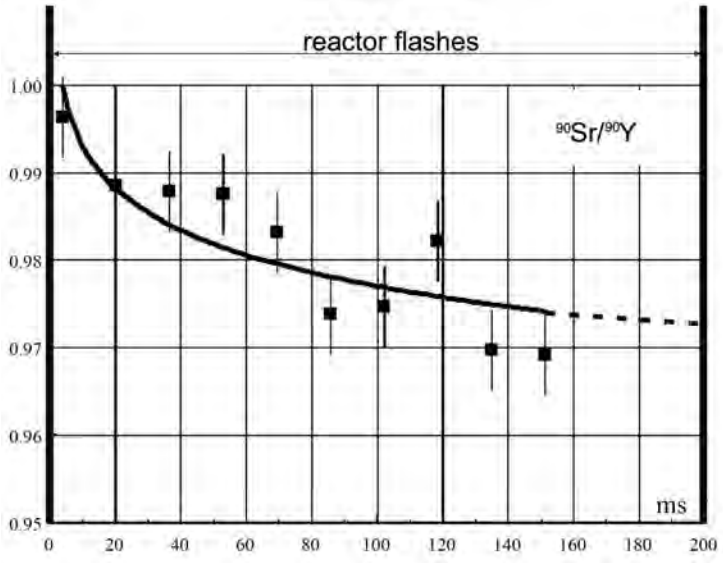


Figure 1. The result of the accumulation of registered beta-electrons emitted by the $^{90}Sr/^{90}Y$ source in the time interval between reactor flashes.

In new experiment, the rate of β-decays in the source ^{63}Ni was measured.

Beta-electrons emitted by ^{63}Ni nuclei were registered using a scintillation detector equipped with a photomultiplier (**Figure 2**).

Using electronics, the photomultiplier pulses were amplified, discriminated in amplitude, and then registered.

In order to protect the sensitive part of the equipment from the penetration of neutrons and γ-quanta, it was placed in a small house made of lead bricks and borinated polyethylene.

The measuring system was located at a distance of about 20 meters from the core of the pulse reactor.

In the absence of the β-source, the measuring system gave false positives no more than once every few minutes when the reactor was running.

When the reactor was switched off, beta decays in the source occurred statistically evenly.

The results of measurements of the effect of reactor neutrinos on the decay rate of the ^{63}Ni isotope are shown in **Figure 3**.

This figure shows a peak of reactor activity, generated approximately 3 ms after the start of data logging. The figure shows that before the reactor flash, the beta-electron count occurs at a certain low level, determined by the established amplitude discrimination.

Without discrimination, the beta-electron count, determined by the activity of the source, is about 6 orders of magnitude higher.

The reactor pulse generates a spike in the rate of decay in the source, followed by the linear decline (on a logarithmic scale, *i.e.* exponential on a linear scale) of the account. This exponent corresponds to a half life of approximately 20 ms, which corresponds to the reference data for the isotope ^{12}B [12].

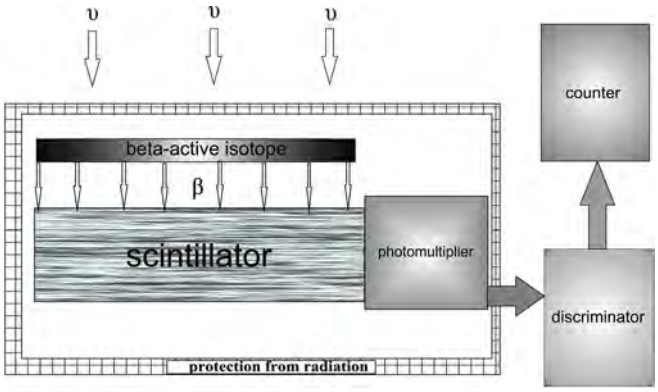


Figure 2. Diagram of the measuring unit.

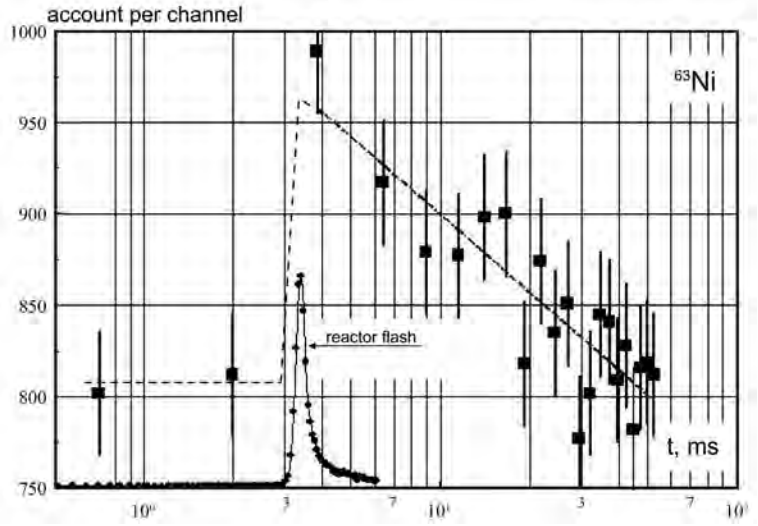


Figure 3. The result of the accumulating registration of beta-electrons emitted by ^{63}Ni. Measurement time is 1 day. The level of amplitude discrimination close to the boundary energy was chosen expermentally. On abscissa: time in ms in the logarithmic scale.

As known the boundary energy of beta-electrons of the ^{63}Ni isotope is only 0.067 MeV, while the boundary energy of beta-electrons of ^{90}Sr/^{90}Y isotope is approximately 2.3 MeV.

When selecting the ^{63}Ni isotope as an object for studying we assumed that the effect of the neutrino flux on an isotope with a lower boundary energy would be stronger. It turned out to be correct.

There seems to be some approximately inverse relationship between the number of beta-electrons generated by reactor neutrinos and their energy.

The physics of this dependence seems to be related to the fact that when neutrino collides with nucleus, which leads to the decay of this nucleus, some part of the neutrino's energy is ultimately transferred to the beta-electron.

And the probability of such energy transfer is greater for small portions of it.

The average energy of a reactor neutrino can be equal to 1 MeV or more.

However, since only a small portion of this energy reaches beta-electron, in

order to register this effect, it is advisable to study it on isotopes with low energy of beta-electrons.

As seen in Figure 3 the decay total account of ^{63}Ni increases under the action of reactor neutrinos by about a quarter per 1 day. The same flux of reactor neutrinos affects the rate of decay of ^{90}Sr/^{90}Y isotopes about ten times weaker (Figure 1).

5. Conclusions

In response to A. Einstein's metaphorical statement, anti-determinists developed quantum mechanical theory of alpha-decay.

The power of the quantum mechanical apparatus and the novelty of this approach have tilted public opinion to the side of anti-determinists.

As a result currently, in the physical community, most believe that radioactive decay is a truly random process.

However, this issue should not be decided by voting.

The way to solve such problems was determined more than 400 years ago by the outstanding scientist of the middle ages W. Gilbert (1544-1603), who formulated the basic principle of natural sciences:

"Theoretical constructions that claim to be scientific must be tested and confirmed experimentally."

The dispute between determinists and anti-determinists about the nature of nuclear decay can be resolved experimentally only.

It seems that the discovered effect of reactor neutrinos on the rate of beta-decay suggests that the nature of this phenomenon was correctly understood by A. Einstein.

Now it seems that the study of the possible influence of reactor neutrinos on the phenomenon of alpha-decay is of great interest.

The author is sincerely grateful to A.V. Strelkov and V. V. Zhuravlev for preparing the high-purity ^{63}Ni beta-source and realization of measurements at the reactor.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Vasiliev, B.V. (2020) *Journal of Modern Physics*, **11**, 91-96. https://doi.org/10.4236/jmp.2020.111005 https://www.scirp.org/Journal/PaperInformation.aspx?PaperID=97775

[2] Falkenberg, E.D. (2001) *Apeiron*, **8**, 32-45. https://pdfs.semanticscholar.org/3a21/346203f836847e60016770d931b324a46be9.pdf

[3] Jenkins, J.H., *et al.* (2008) Evidence for Correlations between Nuclear Decay Rates and Earth-Sun Distance. arXiv:0808.3283v1 [astro-ph]

https://doi.org/10.1016/j.astropartphys.2009.05.004
https://arxiv.org/abs/0808.3283

[4] Vasiliev, B.V. (2017) *Journal of Modern Physics*, **8**, 338-348.
https://doi.org/10.4236/jmp.2017.83023
http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=74443

[5] Vasiliev, B.V. (2015) *Journal of Modern Physics*, **6**, 648-659.
https://doi.org/10.4236/jmp.2015.65071
http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=55921

[6] Vasiliev, B.V. (2019) *Journal of Modern Physics*, **10**, 1487-1497.
https://www.scirp.org/journal/paperinformation.aspx?paperid=96226
https://doi.org/10.4236/jmp.2019.1013098

[7] Vasiliev, B.V. (2017) *Journal of Nuclear Science*, **1**, 2.
http://scifed.com/fulltext/the-electromagnetic-model-of-neutron-and-nature-of-nuclear-forces/21910
https://doi.org/10.23959/sfjns-1000013

[8] Serebrov, A.P. (2005) *Physics-Uspekhi*, **48**, 867. http://ufn.ru/ru/articles/2005/9/a/
https://doi.org/10.1070/PU2005v048n09ABEH003536

[9] Ahmed, S.N., Anthony, A.E., Beier, E.W., *et al.* (2004) SNO Collaboration. arXiv:nuclyex 0309004.

[10] Barnes, V.E., *et al.* (2019) *Applied Radiation and Isotopes*, **149**, 182-199.
https://doi.org/10.1016/j.apradiso.2019.01.027

[11] Hayes, A.C. and Vogel, P. (2016) Reactor Neutrino Spectra. arXiv 1605.02047v1
https://arxiv.org/pdf/1605.02047.pdf
https://doi.org/10.1146/annurev-nucl-102115-044826

[12] Kikoine, I.K., a.o. (1978) Physical Tables, Moscow, Atomizdat. (In Russian)

Journal of Modern Physics, 2020, 11, 616-647
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

The Pioneer Effect: A New Physics with a New Principle

Russell Bagdoo

Saint-Bruno-de-Montarville, Quebec, Canada
Email: rbagdoo@gmail.com, rbagdoo@yahoo.ca

How to cite this paper: Bagdoo, R. (2020) The Pioneer Effect: A New Physics with a New Principle. *Journal of Modern Physics*, **11**, 616-647.
https://doi.org/10.4236/jmp.2020.115041

Received: March 20, 2020
Accepted: May 5, 2020
Published: May 8, 2020

Abstract

Radiometric data from the Pioneer 10/11, Galileo and Ulysses spacecraft indicated an anomalous constant acceleration acting on them, directed toward the Sun, and a gradual growth of the radio signal frequency emitted by the receding transmitter. The reported odd acceleration of Pioneer 10 with a magnitude $\sim 8.5 \times 10^{-10}$ m/s^2 can be explained by an induced gravitational interaction on the S-band signals traveling between the probe and the Earth, arising from the electromagnetic properties of the outer Solar System vacuum zero-point radiation interacting with matter. Their nature is of quantum vacuum origin, and these induced forces act in addition to ordinary gravitational forces, violating the principle of Equivalence. We suggest a new physical theory based on a new principle called "Compensation" as a thinkable explanation for the non-conventional Pioneer effect. The theory of Relation, which is an alternative to the inflationist model, postulates that our universe is made of two antagonistic but complementary structures. The principle of Compensation contradicts Relativity theory, predicts such acceleration and is for the electromagnetic spacetime metric what the principle of Equivalence is for the gravific spacetime metric.

Keywords

Theory of Relation, Principle of Compensation, Two Structures, Negative EM Wavelength of Spacetime, Zero Point Fluctuations

1. Introduction

First published in 1998 (PRL 81, 2858), results from an almost twenty-year study of radiometric data from Pioneer 10/11, Galileo and Ulysses spacecraft have been continuously reported by Anderson *et al.* (1998, 1999, 2002) [1] [2] [3]. Detailed analyses of these data indicated existence of an apparent anomalous,

constant, acceleration acting on the spacecraft, directed toward the Sun [4]. The spacecraft ranging and Doppler velocity measurements are performed by sending a radio signal from the Earth to the spacecraft, from which an active transponder returns it back. The time of the signal travel is used for the position calculation verification and the Doppler effect of frequency change with its high accuracy provides an additional verification of the relative velocity. Considering v_{model} [5], the predicted frequency of the re-transmitted signal observed by a DSN antennae, and v_o the reference frequency. The observed, two-way normal effect can be expressed to first order in v/c as

$$v_{model} = v_o\left[1-2v_{model}(t)/c\right]. \tag{1}$$

The factor 2 is because they use two- and three-way data. v_{model} is the modeled velocity of the spacecraft due to the gravitational and other large forces. This velocity is outwards and hence produces a red shift. The measurements with the use of special programs permitted to observe a steady increment of frequency of the radio signal sent off from the Earth and actively re-transmitted by the spacecraft, *i.e.* on the background of the usual Doppler effect, the drift of the measured frequency toward its growth displayed itself and allowed to conclude that Pioneer 10 slows down. Therefore, the acceleration a_p produces a slight blue shift on top of the larger red shift

$$\left[v_{obs}(t)-v_{model}(t)\right]_{DSN} = -v_o\left(2a_p t\right)/c. \tag{2}$$

v_{obs} is the frequency of the re-transmitted signal observed by a DSN antennae. Anderson *et al.* concluded that there was an anomalous, weak, long-range acceleration of the Pioneer spacecraft in the outer regions of the solar system toward the Sun of $\sim -8.5\times10^{-10}$ m/s^2. The sign indicates that a_p is inward [5]. It shows a steady frequency drift of about -6×10^{-9} Hz/s, or 1.5 Hz, over 8 years (one-way only). This equates to a clock acceleration, $-a_t$, of -2.8×10^{-18} s/s^2. The identity with the apparent Pioneer acceleration is

$$-a_t = a_p/c = -8.5\times10^{-10}\,\mathrm{m/s^2}/c = -2.88\times10^{-18}\,\mathrm{s/s^2}. \tag{3}$$

The S-band carrier radio signal frequency is $v_o \approx 2.29$ GHz

$$-a_t v = v/-t_{cosm} = -a_p v/c \tag{4}$$

$$\begin{aligned}
&\left(-2.88\times10^{-18}\,\mathrm{s/s^2}\right)\times\left(2.29\times10^{9}\,\mathrm{Hz}\right)\\
&=\left(2.29\times10^{9}\,\mathrm{Hz}\right)/\left(-3.472\times10^{17}\,\mathrm{s}\right)\\
&=\left(-8.65\times10^{-10}\,\mathrm{m/s^2}\right)\left(2.29\times10^{9}\,\mathrm{Hz}\right)/c\\
&=-6.6\times10^{-9}\,\mathrm{Hz/s}
\end{aligned}$$

Besides, an independent analysis by Markwardt [6] of radio Doppler tracking data from the Pioneer 10 spacecraft for the time period 1987-1994 confirms the previous observations. The anomalous accelerations observed are constant and persistent. None of the claimed conventional effects can account for the magnitude and direction of the anomaly.

We think that the reported anomalous acceleration of the Pioneer 10 spacecraft of $\sim -8.5 \times 10^{-10}$ m/s^2 (*i.e.* toward the Sun) and the gradual growth of the radio signal frequency emitted by the receding transmitter can be explained by a gravitational interaction on the S-band signals traveling between Pioneer 10 and the Earth, and by the properties of the interplanetary vacuum.

But before coming to this explanation, let's relate what happened around 2011-2012. Until 2011, after many theories had been advanced, there was no satisfactory explanation for this acceleration anomaly affecting the Pioneer 10 and 11 space probes during their transit outside the solar system and measured between 1979 and 2002. An Italian-Portuguese team in 2011 verifies and validates a relatively prosaic explanation of the Pioneer anomaly. The heat produced by radioisotope thermoelectric generators is emitted isotropically (in all directions), but a significant fraction of it is reflected by the back of the large gain satellite dish. The resulting radiation pressure pushes in the opposite direction, *i.e.* in the direction of the Sun. These results have been confirmed by independent teams using thermal modeling software and in 2012 the scientific community believes that the mystery is solved.

We believe that blind confidence in this assumption alone of thermal recoil pressure has made these digital models too schematic. In this respect, we consider it useful to submit here another hypothesis which leads to the opposite conclusion: the thrust is in the direction of displacement of the probe.

First, let's mention the main internally generated heat source comes from the spacecraft electronics box located on the leading side of the spacecraft, but just behind the large parabolic high-gain antenna that points to Earth. The box collects the heat of the various scientific instruments (including the heat of the sun captured by the antenna). Orifices of the box facing the top of the antenna curve evacuate the heat. We assume that the arrival of this heat forms a kind of exhaust gas does not crush on the top of the antenna but dissipates in jets on both sides of the axis of the parabola. These jets slide along the opposite side walls and accompany the curves towards the flared end of the parabolic antenna where they are thrown into the void facing the sun. This is how the probe would be pushed forward, in the direction of travel of the spacecraft, by a force called thrust; this under Newton's third law of motion: each action causes an equal reaction and of opposite direction.

The other source of heat comes from two radio-thermal generators (RTGs), cylindrical in shape, each connected to the main compartment through a truss. These nuclear power sources also emit heat toward the back side of the disc-shaped antenna pointing in the direction of the Earth. When the thermal radiation reaches the main compartment of the equipment, where the majority of the energy is consumed, it joins the gas streams evacuated from the electronic box which travel at high speed to the ends of the horn of the antenna to be expelled and generate the forward thrust at the closed interior apex of the antenna.

We conclude that an additional anomaly with a conventional explanation can

add to the Pioneer effect. And that at the current level of our knowledge of the Pioneer 10 spacecraft and its trajectory, the statistically significant acceleration anomaly still exists and cannot be resolved in standard physics. Its origin must be found in new principles. An unconventional explanation will have astronomical consequences: cosmology, starting from its foundation [7], will have to consider a paradigm shift; standard physics will undergo major renovations, and it will be an opportunity for discovering a new vision of the universe.

In this paper we study the consistency of the data of the Pioneer anomaly with the new theory of Relation. Firstly, in Sec. 2, we introduce the initial conditions of the theory and its new principle. In Part 1 Sec. 3, we show how Pioneer effect is a violation of the equivalence principle. In Sec. 3.1, we postulate the principle of Compensation, which explains the Pioneer anomaly of the outwards redshift in relation with the inward blueshift, and find the deep meaning of this principle in the existence of two structures going in opposite directions. In Sec. 3.2, we suggest that the loss of energy of the photon in the farther interplanetary vacuum induces a gravitational effect like a macroscopic Casimir effect. In 3.3, we understand the vacuum by a possible effect from unknown physics: a gravitational frequency shift of the radio signals proportional to the distance to the spacecraft and a loss of density energy in the intermediate medium. We argue that the gravitational interaction of the S-band radio with a growing empty vacuum might be responsible of the anomalous acceleration of the Pioneer spacecraft. In 3.4, according to the CP, we construe bond between the gravitational photon of the structure of condensation and the cosmological photon of the structure of expansion: energy lost by the latter is retrieved by the former. We discern, in Sec.3.5, gravific spacetime and EM spacetime, and explain in 3.6 the relation between the acceleration of Pioneer inward and its deceleration outwards. Predictions in Sec. 3.7. In Part 11 Sec. 4, the unaccepted concept of "tired light" seems consistent with data of the anomaly and is integrated within the framework of the theory of Relation which claims, in 4.1, that gravitation and electromagnetism are two sides of one force. In Sec. 4.2, with the intention of dispelling a huge discrepancy and to be reconciled with anomalous data, we propose a negative cosmological constant with a varying density of energy through time, which implies a positive pressure. In Sec. 4.3, we discuss of links between CP, zero point fluctuations, induction of gravity and inertia. Along a different line in 4.4, we concentrate on connections between theory of Relation and quantum gravity through Pioneer effect. With correlations and assumptions concerning the graviton, we write down some characterizations of a quantum gravity to be constructed, finally, the conclusion in Sec. 5.

2. Relativity, Theory of Relation and Principle of Compensation

Special Relativity (SR) concerns isolated system, free of outside influence, in a completely empty and homogeneous space in which the speed of light is an ab-

solute constant. General Relativity (GR) is SR looking at the influence of gravitation. It is based on the principle of Equivalence (EP) considering the equality everywhere between gravitational mass and inertial mass. Matter, which is gravity, curves spacetime.

The new theory called "Theory of Relation" regards that not only matter but also vacuum interact with the system. Vacuum is not empty space but ocean of interacting negative and positive energies. The reported anomalous acceleration of the Pioneer 10 spacecraft is a consequence of this misunderstood part of our physics. This theory gives an interaction between the electromagnetic (EM) spacetime of expansion and the gravitational matter of condensation. This new cosmological model initially assumed that

1) Our universe is made of two complementary and interpenetrated structures, one for condensation with a gravific spacetime and EM matter (Einstein), the other for expansion with a flat EM spacetime and ordinary matter (Lorentz-Maxwell).

2) Since the big bang, the EM structure of expansion is decreasing giving up its energy to the positive gravitational increasing structure of condensation. A perpetual annihilation of the negative energy-mass is transformed into a continual creation of positive energy-mass. The first structure of condensation represents the positive solution of Dirac's equation of energy while the second structure of expansion express its negative energy solution which was eliminated by a mathematical trick [8].

3) The EM wave of spacetime is supported by an inhomogeneous vacuum filled of "minimal" negative energy perpetually in interaction with positive matter. The speed of light is a constant but there is an anisotropy making it not absolute.

4) Ether exists. There is no precise definition for it and it is a lumber room for scientists. In our theory, it is close to the Lorentz's ether theory: a negative EM spacetime wave in a quantum void, which is a *plenum*.

5) Electromagnetism in macrocosm is the negative side of gravitation [8] [9]. If gravitation is the only force between planets in the structure of condensation, electromagnetism is the driving force of the structure of expansion. Gravitation of the first structure is "positive" relatively to the "negative" electromagnetism of the second structure.

6) The theory is in a framework of "cyclic universe theory". Our model combines the cosmology of the big bang with the continual creation of matter. Our big bang starts with the big crunch of a pre-big bang universe. This pre-compression resolves the problems of causality, conservation, flatness, horizon, without being committed to the cosmological standard model of inflation, strings, higher dimensions, or to branes and orbifolds. The concepts that the universe "comes back full of negative energy-mass", and the "normalization" of the negative energy solution of the classic equations, are the central ideas of this cyclic model.

Our complex universe is dual: positive and negative. The negative part of the

universe, which is a universe by itself, is disintegrating, "creating" our actual positive universe. The compensation principle (CP) asserts that the permanent loss of negative energy of the expanding EM wavelength of spacetime induces the positive gravific spacetime matter. Flat EM spacetime can yield induced gravity to ordinary matter. Gravific spacetime matter produced by the expansion can flatten the EM spacetime.

Note on matter and antimatter: In the theory of Relation, our dual universe contains two opposite structures of energy-mass, each has two roots for both positive and negative energy, each containing its own particles and antiparticles, being dependant of Dirac's equation on energy

$$\pm E = \left(m^2c^4 + p^2c^2\right)^{1/2}. \tag{5}$$

In each structure, antimatter is matter that is composed of the antiparticles of those that constitute normal matter. In a larger sense, antimatter is the extension of the structure of expansion, while our normal matter constitutes almost entirely the structure of condensation. This asymmetry of matter and antimatter would come from a kind of continual transmutation attributed to the disintegration, all along the cosmic expansion, of the negative antimatter-anti energy into the creation of positive matter-energy of the observable universe.

PART 1

TWO STRUCTURES AND TWO PRINCIPLES

3. Violation of the Principle of Equivalence in Empty Inertial Space

In empty space of GR, a supposed observer inside the Pioneer spacecraft will theoretically undergo a force equal to "inertial mass × acceleration", also equivalent to a force equal to "gravitational mass × intensity of the gravitational field" [10]

$$F_i = M_i \times a = F_g = M_g \times g. \tag{6}$$

This equivalence is corroborated by experiences on Earth. But observation on Pioneer gives

$$\left(F_i = M_i \times a\right) \neq \left(F_g = M_g \times g\right). \tag{7}$$

$$\text{Precisely} \quad \left(F_i = M_i \times a\right) < \left(F_g = M_g \times g\right). \tag{8}$$

Keeping the equality between the masses, *a* decreases while *g* increases. Anomalous deceleration outwards is interpreted like anomalous acceleration toward the Sun [5]

$$\Delta g = -a_i c = a_p = -8.5 \times 10^{-10}. \tag{9}$$

This anomalous acceleration, which signifies an extra measurable gravitational potential, counteracts the EP saying "In a freely falling local frame (even including self-gravitating bodies), no influence is measurable of the gravitational potential which may cause an acceleration of the frame as a whole" [11].

Radio-signal interval is shorter than expected meaning linear radius (Newtonian) is shorter (range). With GR, shorter distance wants to say a greater curvature due to an excess of sunward gravity [12] [13]. Let us suppose a clock situated on a probe in a fieldless region of space. With GR, if we take the origin of the space coordinates in the center of an atom, we can put $\xi(x)=\eta(y)=\zeta(z)=0$, and have

$$s^2=-c^2T^2=g_{44}t^2$$

$$t=Tc/\left(-g_{44}\right)^{1/2}. \tag{10}$$

It is only in a fieldless space that we have $g_{44}=-c^2$, thus $t=T$.

Suppose a clock situated on Pioneer 10 in the same "fieldless" region of space. An anomalous acceleration is observed: we say that a downward trend of minimal energy with distance in the vacuum induces gravity. So the gravitational field g_{44} in this false fieldless is different from $-c^2$, say $g_{44}=-c^2(1-\gamma)$. Thus the time of vibration is altered to

$$t=T/(1-\gamma)^{1/2} \tag{11}$$

or, as the deviation γ is small, approximately

$$t=T/(1+\gamma/2);\ \left[\gamma=2\Phi/c^2=2GM/Rc^2=2v^2/c^2\right]. \tag{12}$$

For GR and the EP, a clock in the vacuum undergoes no gravity. The CP says there is a difference in the beating of the same clock for which the difference of the induced gravitational field in the vacuum given by g_{44} has the relative value γ.

Pioneer effect is a clear violation of EP, an anomaly in the redshift of GR. It is as if the loss of density in the vacuum was turned into gravitational energy, like if a negative energy was converted into a positive energy. Energy of both systems is *conserved* if what constitutes a non-standard channel of energy-loss by *a* is recuperated by *g*. A new principle is necessary to explain this anomalous effect and maintain this conservation.

3.1. Principle of Compensation in Empty EM Space

The principle of compensation says that the decrement of negative EM energy-mass during the expansion induces a proportional and opposite increment of the positive gravitational energy-mass. Considering that further distance from matter lowered minimal density energy in vacuum, assuming that a long-range negative scalar field interacts with matter inducing gravity, the CP permits the transformation of negative EM energy from the increasing wavelength of space-time into positive gravitational matter. According to this principle, the negative EM energy of vacuum slowly lessens with distance inducing the lost into positive gravitational energy. The progressive deceleration of the space probe outward, interpreted as a constant sunward acceleration, is related to the deep general idea that the scalar field of the vacuum with an appropriate potential of negative energy is continuously able to transform its energy into a positive

gravitational energy. It explains the anomaly of the outwards redshift in relation with the inward blueshift.

In a *general* sense, the EM wavelength of spacetime is assimilated with the cosmological constant interpreted as dark energy and corresponding in our theory to the negative mass-energy containing the baryons at speed under c and the bosons at speed *c*. In a *special* sense, the EM wavelength of spacetime is associated with the energy and with the bosons at the speed of light.

The situation of the cosmological photons streaming from this *special* EM wave of spacetime, related by CP to the gravitational photon of the structure of condensation, is illustrated in a sort of double vacuum wells which forms a triangle ◄ with A at left, E down and G up. A has a double role: it is the EM charge source for the EM cosmological photon on the structure of expansion (similar to the metric of special relativity) and it is the attractive mass for the gravitational photon on the structure of condensation (similar to the metric of GR). The branch AG is of a gravitational well. The branch AE belongs to an empty quantum well. Both wells are different but related.

When cosmological EM photons in space fall from A to E into the steady slowly growing vacuum which forms a "quantum well", they lose energy (redshift). The energy lost at point E induces energy in gravitational photons on the structure of condensation at point G falling down toward A into a "matter well". So the gravitational photons rolling down the potential well will have an extra energy (blueshift) coming from the vacuum energy, adding the kinetic energy to a standard gravitational field obeying the EP. We can also say that the EM photon tumbling from A to E into a vacuum well becomes less energetic and slightly colder by $(v/c)T^o$, where the inferred velocity is $v < c$. The gravitational photon at G recovers the energy lost by the opposite photon E, is hotter by $(v/c)T^o$, and is more energetic before he falls into this kind of gravitational well than the conventional gravitational photon gaining energy like a ball rolling down a hill.

When cosmological photons climb back out from E to A, they are blueshifted, they gain EM energy. On the contrary, gravitational photons rolling up a matter well lose energy, are redshifted. It is like if the lost gravitational energy was transformed into EM energy. The deep meaning of CP is that when there is less EM mass-charge repulsive force in the structure of expansion going forward with the arrow of cosmological time, there is more mass-matter attractive force in the other structure. When there is more EM energy in the structure of expansion going backward against the arrow of cosmological time, there is more negative EM energy-mass, more quantum gravity, less classical gravitational energy-mass-matter, an antigravity. The conservation of mass-energy law is basically respected because there are two structures with two incompatible energies: the negative EM energy and the positive energy of rest mass. CP allowed their conversion in each other. The annihilation of one form of matter-energy is the creation of the other. The present universal current is the disintegration of the negative EM energy-mass converted into the positive gravitational energy-mass.

With the theory of Relation, the cosmological photon is bound up with the gravitational photon. Thus, in the same proportion that the cosmological photon will be stretched and redshifted to a longer wavelength in the deepening potential well of the fabric of spacetime itself, the gravitational photon will be contracted and blueshifted to a shorter wavelength in a gravific potential well. This paradoxical situation gives the connexion between the compensation and the equivalence principles. Note that inflation or cyclic brane cosmology models, with the dark energy causing universal expansion to accelerate toward a big rip, are pure violation of EP.

3.2. Macroscopic Casimir Effect: Induction of Gravity

Photons of the interplanetary vacuum simply lose energy by traveling through space toward the infinity. In deep space there is an increase of wavelength of photon proportionally to the time/distance while passing through the space undisturbed [14]. The loss of energy induces a gravitational effect like a macroscopic Casimir effect [8] [9]. Gravity is induced when the EM energy of space is lowered between two plates. It implies a mechanical action which is equivalent to the *ad hoc* explanation saying that shorter external waves pushed plates one toward the other. It seems likely that redshifts may be due to a decelerating expanding EM spacetime transforming its loss of energy into the condensed matter of the gravitational structure. The observed excess redshift express both a steady decrease of the minimal energy of vacuum proportional to the distance and an induced gravity of matter. The observed excess blueshift express both a shorter distance and a higher frequency from positive quanta of energy.

Quantum vacuum field induced a constant gravity in fieldless space which is like the action of a constant gravitational field such as occurs at the immediate surface of a celestial body which, in accordance with GR, can be replaced by an acceleration of the observer with the same value g directed opposite to the attraction. Pioneer effect is related to energy. Induced gravity by quantum vacuum aimed at the Sun means mass enlargement and lower gravitational potential (negative gravitational potential energy becomes smaller when bodies come closer) [15]. This anomalous lower potential energy from Sun is amount to an increase of mass (if c is the speed of gravity and light, and if the distance is anomalously shorter, M_s becomes more massive in $GM_s/c^2 = r$). In fact, the CP is the underlying mechanism bringing a perpetual creation of positive energy-matter. Does it say there is no more conservation of energy and mass? Absolutely not, but this law must be enlarged to the notion of negative energy and mass in agreement with $\pm E = \pm\left(mc^2\right)$.

3.3. Regarding a Lower Density of the Vacuum

Anderson *et al.* [5] reported D.F.Crawford [16] suggesting a possible effect from applications of known physics: a gravitational frequency shift of the radio signals that is proportional to the distance to the spacecraft and the density of dust in

the intermediate medium. In particular, he has argued that the gravitational interaction of the S-band radio signals with the interplanetary dust may be responsible for producing an anomalous acceleration similar to that seen by the Pioneer spacecraft. The effect of this interaction is a frequency shift that is proportional to the distance and the square root of the density of the medium in which it travels. Didon, Perchoux and Courten [17] proposed similarly that the effect comes from resistance of the spacecraft antennae as they transverse the interplanetary dust. This is related to more general ideas that an asteroid or comet belt, with its associated dust, might cause the effect by gravitational interactions or resistance to dust particles [18]. "However, these ideas have problems with known properties of the interplanetary medium: infrared observations rule out more than 0.3 Earth mass from Kuiper Belt dust in the trans-Neptunian region. Ulysses and Galileo measurements in the inner solar system find very few dust grains in the 10^{-18} - 10^{-12} kg·m^{-3} range. The density varies greatly, up and down, within the belt (which precludes a constant force) and, in any event, the density is not large enough to produce a gravitational acceleration on the order of a_p [5].

We think that the effect does not come from resistance of the spacecraft antennae as they transverse the interplanetary dust but on the contrary from a void of density. This is related to the more general idea that emptiness might cause the effect by gravitational interactions. It is not the resistance to "positive" dust particles but their absence which is generating gravity. Induced gravity in Casimir effect is not caused by presence but by absence of energy.

We suggest a possible effect from applications of unknown physics: a gravitational frequency shift of the radio signals that is proportional to the distance to the spacecraft and the loss of density energy (dust, etc.) in the intermediate medium. In particular, we argue that the gravitational interaction of the S-band radio signals with the growing empty space of the interplanetary vacuum may be responsible for producing drag acceleration similar to that seen by the Pioneer spacecraft. The effect of this interaction could be a frequency shift that is proportional to the distance and the square root of the "lack" of density of the medium in which it travels (not, as expected, proportional to the square root of the density of the interplanetary dust) [16].

It is assumed that the density of the vacuum $p_{vacuum} = p_{solar\ system}/n$ differs from the density in solar system (ecliptic) significantly. Index $n<1$ is the averaged characteristic of the negative energy layer thickness of which decreases by the law l (distance) = vt (outwards) between the probe-transmitter and the Earth-receiver [19]. The Pioneer effect comes from resistance of the spacecraft antennae as they go through the interplanetary medium. This resistance, like associated gelatine, which is not real dust particles, is the negative pressure of a negative energy, which is the scalar medium of the vacuum. It induces a gravity which has a real mechanical effect similar to the Casimir effect. It is also presumed by the CP that this index (n toward 1) is likewise the averaged cha-

racteristic of the positive gravitational energy layer thickness of which increases by the law $l = vt$ (inward).

Having said this, we do not exclude the presence of dark matter as associated gelatine forming the resistance. Without elaborating on the nature of dark matter, it is in the theory of Relation an intermediate state between negative and positive energy-matter. The creation of this exotic matter is always accompanied with a creation of an exotic space, and both produces a *gravitational flux density*, both concerns non-baryons and non-bosons originating from the structure of expansion on a very slow way to become baryons and bosons in the structure of condensation. The specific case of Pioneer relates to bosons and "dark space" (void).

3.4. Gravitational and Cosmological Photons

The measurements allowed observing an increment of frequency of the radio signal sent off from the Earth and retransmitted by Pioneer 10. Light of frequency v can be regarded, according to quantum theory, as consisting of quanta of energy $E = hv$. These have an inertial mass $m = E/c^2 = hv/c^2$ and this is, according to the EP, equal to its gravitational mass. When light quanta hc have travelled the distance l toward the field of gravitation g, their energy has increased by gl/m. The receiver on Earth detects a frequency

$$hv' = hv + g/hv/c^2 . \tag{13}$$

According to the CP in the deep space, we must add an extra induced gravity links up to vacuum. The gravitational interaction is caused by the focusing of the signal photons in curved space; in this case, the curvature is related by an anomalous induced gravity characterized by the *lack* of density of the interplanetary *positive* dust. When light quanta hv have travelled the distance l toward the field of gravitation g, their energy has increased by

$g/m + \Delta g/m$. Hence at the end of this journey the energy of quantum is:

$$hv'' = hv + \left(g/hv/c^2 + \Delta g/hv/c^2\right). \tag{14}$$

The receiver on Earth detects a frequency slightly higher than the one detected according to the EP ($hv' = hv + g/hv/c^2$). It is a blueshift for the "gravitational" photons on the structure of condensation. Radio-signals intervals are shorter. The spacecraft in deep space undergoes more frequency which is not increment of speed because there is no recession: data proved Pioneer 10 is slowing down going outwards. On the standpoint of the spacecraft going outwards, the frequency observed on Pioneer (v'') is anomalously more diminished than the expected frequency (v').

$$v'' = v' - \Delta v' = v\left[1 - \left(v + \Delta v\right)^2/c^2\right] = v\left[1 - \left(g + \Delta g\right)l/c^2\right]. \tag{15}$$

The observed redshift is greater than the calculated one. So the extra g from gl/c^2 is an outwards deceleration interpreted, with a change of sign, like acceleration toward Earth-Sun, which is an application of CP.

What about "cosmological" photons on the expansion structure? In deep space they follow the EM expansionist arrow of Time. Their sources are the negative matter-energy at the beginning of our cosmological time. They are in a position of a gravitational photon (emitted by star or planet) going against the field of gravitation g and losing energy with outwards distance. With regard to the EP and Doppler effect, when light quanta hv have travelled the distance l with an expansionist EM field, their energy has decreased by g/m giving the frequency $hv' = hv - g/hv/c^2$. In empty space, the photon loses also an additional energy because a downward trend of minimal energy with distance. Hence at the end of this journey, according to the CP, there is an anomalous redshift and the receiver on Earth should detect a frequency slightly lower than the one expected [12] [13]

$$hv'' = hv - \left(g/hv/c^2 + \Delta g/hv/c^2\right). \tag{16}$$

Another way to say this is that in the deep space out of our solar system the probe follows an increasing EM spacetime wave. We postulate that with distance in vacuum the minimal negative EM energy of expansion is slowly diminished. It is a real effect in addition to the usual "Doppler effect". There is an anomalous loss of energy in the vacuum against the EM source giving an anomalous redshift. The energy of a quantum alongside with Pioneer going outwards is only

$$hv'' = hv' - \Delta hv' = hv\left[1 - \left(v + \Delta v\right)^2/c^2\right] = hv\left[1 - \left(g + \Delta g\right)l/c^2\right]. \tag{17}$$

This equation without h is like (15), where an anomalous acceleration toward the gravitational source gives the same redshift. It is like if the EM energy was converted into gravitational resistance or acceleration toward matter. Pioneer 10 in the deep space of expansion slows down going outwards; the cosmological photon loses energy, which fuels the gravitational potential. According to the CP, the energy lost by the cosmological photon ($-g/hv/c^2$) is retrieved by the gravitational photon ($+g/hv/c^2$). The observed redshift on Earth is higher than predicted, but the real effect is a blueshift. In the vacuum, "lost energy" is equivalent to "gain of gravity", "anomalous redshift" from light quanta is equivalent to "anomalous blueshift" from anomalous acceleration (induced gravity).

3.5. About Time and Metric

Theory of Relation discerns two real spacetimes: the gravific spacetime, or Einsteinian spacetime, associated with positive energy and condensation; and the EM spacetime, or Lorentz-Maxwell spacetime, associated with negative energy and expansion. And two distinctive times going from past to future: cosmic time for the gravific spacetime (cosmic time of GR); and the cosmological time for the EM spacetime. Cosmologic time goes along with macroscopic thermodynamic time (entropy). Considering the two structures, electromagnetism becomes the counterpart of gravitation in the macrocosm [20] [21].

Usually we consider cosmic time, atomic time and ephemeris time. Cosmic

time is the time of GR: dilatation of time when space contracts with greater curvature. Atomic time is also called "proper time". Ephemeris time is associated with Kepler-Newton time. It is the uniform time of Solar system; planets around the Sun. In a standard GR where empty space is without energy there is an agreement between those times. Non-conventional Pioneer effect contradicts relativity theory.

There is equality between atomic time of DNS antenna, cosmic time, ephemeris time but not with cosmological time on which they are depending. EM time (link with absolute time of our universe) is still ignored because not only the positive energy but also the negative energy of the vacuum is not considered. The origin of the anomalous signals is related to this cosmologic EM time. Cosmological time of a clock in the scalar field of vacuum (negative scalar) is longer, because the EM wave of spacetime is longer with the distance and loses some negative energy (tired light).

Atomic time of a clock on probe Pioneer in vacuum is also longer because matter gains the same negative energy converted into positive energy. Atomic time of DNS antenna takes the coincidental point-time of a time clock accelerating toward Sun (blueshift) corresponding to an anomalous decelerating Pioneer outwards (redshift). This shorter atomic time is not the same as the longer "proper time" of the atoms of Pioneer in deep space. In this regard, there is two "atomic times".

The radio signal sent off from the Earth and actively retransmitted by the spacecraft constitutes *a* "*clock*" *which is not in a constant relation with an atomic clock on Earth*. Perturbed shorter intervals affect the "temporal length" of the wavelength of the spectral narrow color lines even if it does not change the type of atom emitting the light. This is a violation of the relativistic ds^2 based on Minkowski metric in space which embodies an imaginary time and directly no EM field. Our interpretation is that "temporal length" of the atom depends of distance and of a Lorentz-Maxwell metric with an EM wave of spacetime producing the cosmological time of our universe.

Note that, on Earth, only the atomic clock in relation with Pioneer going outwards with an anomalous deceleration is affected by the change of frequency. It reflects an anomalous and constant modification of the state of the vacuum surrounding Pioneer. The usual atomic clocks on Earth and on the planets are influenced just by the physical or mechanical local conditions in the near limited domain in which is this clock [22]. The irregular deceleration outwards interpreted as an anomalous acceleration inward is not seen on planets and there is no perceived change in the frequency of the usual atomic clocks, because orbital planets are in a *finite and stationary movement*.

This being said, ephemeris time apparently does not change. But beyond Neptune, Pioneer 10/11 radiometric data indicated a constant skewing between predicted and observed Doppler shifts [23]. It is a blueshift: radial distance become shorter because an excess of gravity makes the range of Newtonian radius

shorter (greater curve with GR). The time of ephemeris, which is Newtonian time, is shorter and so is the range disturbance of the proportional radius.

3.6. Relation (Not Coincidence), Ether and Two Structures

There is a relation between, on one hand, the anomalous acceleration of the probe toward the Sun ($-a_p$) and the acceleration of the clock ($a_t = a_p/c$), and, on the other hand, an anomalous decrease in acceleration outwards and the increase of EM wavelength in deep space.

We have postulated the existence of two structures going from past to future but oppositely directed in space: condensation with the positive energy of gravity and expansion with the negative energy of electromagnetism. In this respect, EM force on cosmic scale is a counterpart comparable to gravitation.

GR is a local structure of condensation with space as a gravific ether and matter as an EM field. Gravitation is associated by GR to properties of spacetime. With Einstein, the field of electromagnetism in space is not determined by the gravific ether and its links in the ether are secondary, almost nonexistent [24]. SR associated Maxwell's electromagnetism to properties of spacetime and, in our view, this association constitutes the second structure, the one of the global expansion, initiated by Lorentz-Maxwell with an EM ether field as space and the ordinary matter. This second structure, inspired by Dirac, is a negative ocean where the negative energy-matter is perpetually transformed into the positive matter-energy of our world. Searching a physical meaning for the negative energy in his equations, Dirac uses above all the passage of positive energy toward the ocean of negative energy. Relation theory considers the ocean of negative matter-energy as the matrix of our world and its "evaporation" into the space-time-matter of our positive world follows the recessional motion with the arrow of time from past to future. Both structures exist and are necessary to our universe but properties of each structure are of different nature. The so-called greatest blunder of Einstein was not about the cosmological constant but the denial of a second structure in our world. It was more a limitation curbing unification than a mistake, since "ether" was set aside.

We assume that Pioneer follows the Lorentz-Maxwell structure in which the arrow of cosmological and thermodynamic times is going outwards toward expansion and entropy. It gives a deceleration of clock time, because longer cosmological time corresponds to a deceleration outwards and an increase of EM wavelength in deep space. Pioneer going outwards was under steady lower frequency drift of about 6×10^{-9} Hz/s which equates to a clock deceleration of 2.8×10^{-18} s/s^2. The identity with Pioneer deceleration in deep space is $+a_t = +a_p/c$ (sign + means a_p is outwards). This corresponds to the observed anomaly using the Doppler effect: the redshift is more than the calculated and verified by the frequent position measurements using the time of travel of the radio signal.

In fact, the location and the motion of the space vehicles point out less velocity [14]. Analysis of Pioneer's radio-metric tracking data has consistently indi-

cated at heliocentric distances the presence of a small Doppler frequency drift which is a blueshift. Anderson *et al* interpreted it as a constant sunward acceleration of spacecraft [25].

The radio-signal known as the Pioneer anomaly follows the structure of Einstein-Newton with the positive energy of gravity. This structure of condensation and the cosmological time are going toward the future with the same arrow of time (→), but both structures have opposite *spatial* direction. Cosmic time follows the spatial direction of condensation (←). In this respect, let's say that our universe is a spacetime with a spatio-temporal dynamics. Spacetime of the expansion structure is *temporally* flat (not completely) with a positive or negative curvature, while the spacetime of the condensation structure is *spatially* curve.

Pioneer going sunward under steady higher frequency drift of about -6×10^{-9} Hz/s which equates to a clock acceleration $-a_t = -a_p/c = -2.8\times10^{-18}\,\mathrm{s/s^2}$ (sign means inward). The acceleration of Pioneer, the acceleration of clock and the frequency are higher. It's a blueshift.

There is an interpenetration of both structures giving opposite results and connected by the CP. Expansion is the dominant structure with cosmological EM time and negative energy begetting the condensation structure with gravific space and positive energy. Let's take $-a_t = +a_p/c$ (negative and positive signs mean inward and outwards). Pioneer going outwards slows down by following the way of spacetime EM in growth. The inverse of a growing universal time gives $-a_t$ which is a lower drift, or lower clock acceleration, sunward† (see 3.7).

The underlying mechanism is that the energy of the negative EM spacetime wave of expansion loses of its energy with the distance so generating the positive energy of the gravity. Longer cosmological time gives a diminishing frequency with distance because the negative energy density of the vacuum scalar field is slowly decreasing by generating an induced gravity, like a Casimir effect. The negative energy lost by space wave is converted into positive energy or, in other words, quantum vacuum is inducing gravity. Equivalently, it gives an anomalous acceleration of clock time, a blueshift going inward. Blueshift means a shorter wave of cosmic time on the condensation structure. The velocity of Pioneer outwards slows down; the distance and the time of travel of the radio signal are shorter.

This relation is not coincidence but an obliged natural link between both structures, an everlasting exchange between negative and positive energies. They are related by the CP: lost energy of expansion is recuperated into positive energy of condensation. There is more gravity-energy in condensation, because with distance there is less energy in the EM wave of expansion. This anomalous gravity-energy is equivalent to the anomalous shorter distance. More than a simple coincidence, $-a_t = +a_p/c$ express a perpetual relation between condensation and expansion, between both energies.

3.7. Predictions

The observations made by WMAP, as well as the high redshift supernova data, are usually explained by an accelerating expansion of the present universe. However, with the current quality of the supernova 1a data, the allowed parameter space is wide enough to accommodate the decelerating models as well [26]. Moreover, the satellite XMM-Newton of Agency's European Space X-ray observatory (ESA) has returned data about the nature of the universe indicating that the universe must be a high-density environment, in clear contradiction to the "concordance model" relying on the theory of inflation. In a survey of distant clusters of galaxies, the results of XMM-Newton revealed that today's clusters of galaxies are superior to those present in the universe around seven thousand million years ago. Such a measure also goes toward a decelerated expansion [27].

At this stage we can wonder if we do not make wrong road from the initial conditions, if we really understand the relation between masses and gravitation, if the official accreditation in 1998 of a positive cosmological constant from the observations of type 1a supernova was not premature and misinterpreted. We foresee a rectification in the future. We think, for reasons else than those required by the classic big bang and by the decelerating Einstein-de Sitter model, which used to be the favoured model before these observations, that the global expansion decelerates since the beginning. The deceleration of Pioneer in the deep space where the wavelength of spacetime grows is in itself an experimental proof of a decelerated world.

In the future, there will be inward a *predictable* less acceleration for the probe, less clock acceleration, less blueshift, less shorter time interval between signal pulsations[†] (see 3.6). In other words, with the distance and the time outwards the anomalous acceleration inward will be weaker, unless there is a sudden attractive body or an unexpected pocket of lower vacuum level. A longer cosmological time corresponds to a declining anomalous acceleration toward Sun. What we call an acceleration of time toward Sun is a deceleration of time toward infinity; the acceleration of Pioneer toward Sun is a deceleration outwards, being astride upon a slowed down expansion associated with the EM spacetime wave.

Notice that a number of people have pointed out

$$a_H = cH \to a_p = a_t c \to \sim 8\times 10^{-10}\ \mathrm{m/s^2} \tag{18}$$

depending on the value of H (inverse of the Universal Time) or a_p [5]. The use of the time of travel of the radio signal gives an additional redshift going outwards: lower frequency, longer wavelength. It showed that there is here no Hubble effect (recession: faster speed of the galaxies with distance) corresponding to Doppler effect. They are prone to conclude that there is no recession at all and no expanding universe because the Hubble redshift is distance caused.

Our view is that the redshift is linked with the EM wave of spacetime of the second structure. A longer spacetime wavelength, or "tired light" in immediate deep space, points out a deceleration of expansion going toward the future.

There is no contradiction with Hubble discovered nebular redshift versus distance explained through recession as long as it is toward the past on the structure of cosmological time: to see weakly the distant galaxies implicates see earlier in the past and a faster speed toward the point of origin. This means that the universe does not accelerate toward the future, and that its speed at the beginning was almost the speed of light.

PART 11

TIRED LIGHT MECHANISM: THE ZPF THEORY AND THE PRINCIPLE OF COMPENSATION

4. Tired Light Integrated into Theory of Relation

Till now, relativity and Doppler effects went along very well to explain the shifting of spectral lines. When astronomers look at the spectrum of light from a star or galaxy, they see the precise wavelength of the narrow, colored spectral line telling scientists what type of atom has emitted light. The Doppler effect is the apparent change in the wavelength of sound or light caused by motion of the source, the observer, or both. Waves emitted by a moving object as received by an observer will be blueshifted if approaching, redshifted if receding. The change in frequency depends on the speed at which the object moves to or away from the receiver [24]. With relativity, if we suppose two stars S_1 and S_2 link up by a rigid stick, a monochromatic light sent from S_1 to S_2 and reflected from S_2 toward S_1 would have the same frequency than the one measured by a clock in S_1: the number of wavelengths along the stick will be the same.

A radio signal sent from Earth to Pioneer 10 and retransmitted to Earth would not have the same frequency as that initially measured by a clock on Earth: the observation of an abnormal blueshift indicates a wavelength that disappeared along the supposed stick between them. On the other hand, there is a cosmological redshift which excludes the Hubble redshift with its distance\source proportionality, interpreted as the recession of galaxies. With distance, the light around the probe seems as tired as its speed.

Pioneer effect is the first support to explain it by the concept of "tired light". One of the difficulties before was that there was no consistent mechanism known that could give a tired light effect: it was not related to transmission in a transparent medium. Data from Pioneer 10 give evidence that there is plasma of space slowing down the photons.

Tired light is a class of hypothetical redshift mechanisms that were proposed as an alternative explanation for the redshift-distance relationship. It was first proposed in 1929 by Fritz Zwicky [28] who suggested that photons might slowly lose energy as they travel vast distances through a static universe by interaction with matter or other photons, or by some novel physical mechanism. Since a decline of energy corresponds to an increase in light's wavelength, this effect would produce a redshift in spectral lines that increase proportionally with the distance of the source. The term "tired light" was coined by Richard Tolman in the early

1930s [29]. Today this model is considered highly speculative even if it was a serious alternative intended to replace the competing cosmologies of the big bang and the steady state, which both assumed that Hubble's law is associated with the metric of expanding space.

In the theory of Relation, tired light becomes a mixture of big bang and steady state cosmologies. The redshift is due to recessional motion and the photons emitted by a nebula lose energy on their journey to the observer. Pioneer 10 going outwards loses some frequency. The Pioneer effect is important because it demonstrates that light in space cannot be explained purely as a wave phenomenon. Relativity and the Maxwellian theory of EM wave cannot explain any real scattered energy and any shift in wavelength. Light must behave as a stream of particles whose energy is proportional to the frequency in order to explain what a Compton scattering for the space seems. This dispersion of energy, like a Compton effect in space, implies discontinuity in space instead of continuity. If photons can impair energy through some interaction with the medium (a plasma) located between the source and the observer, it requires a different metric than the one used by relativity [30].

4.1. Gravitation and Electromagnetism: Two Sides of One Force

The EM force in the vacuum, shown above, arises from a proposal of Sakharov [31] based on the work of Zeldovich [32] in which a connection is drawn between Hilbert-Einstein action and the quantum vacuum [33]. This leads to a view of gravity as "a metric elasticity of space". There were some early pioneering attempts, inspired by Sakharov's conjecture, to relate gravity to the vacuum from a quantum-field-theoretical viewpoint. The first step in developing Sakharov's conjecture in any detail was the work of Puthoff [34]. Gravity is treated as a residual force in the manner of van der Waals forces. Unlike the Coulomb force, which, classically viewed, acts directly between charged particles, this interaction is mediated by extremely tiny propagating secondary fields created by the zero-point fluctuations (ZPF)-driven oscillations [35], and so is enormously weaker than the Coulomb force. Puthoff's attempt to derive a Newtonian inverse square force of gravity proves to be unsuccessful.

Our approach, based on the theory of Relation, considers our whole universe as a white hole estimated as around 10^{53} kg. The radius for the positive energy-mass ($GM^o/c^2 = GE/c^4 = R_U$), as long as it is a Schwarzschild radius, is also a gravific wave $t_o c$. The same radius is the cosmological EM wave of the negative-energy mass of the vacuum ($t_o c_{EM} = h/mc = hc/mc^2$). This means that during the expansion, triggered by the big bang, the negative mass-energy of a negative universe turned into a positive universe with a positive mass-energy. Our universe would be part of a cycle universe. Dark matter would be a negative mass-energy waiting to be transformed into positive mass-energy.

In classical physics, empty spacetime is called the vacuum. In quantum me-

chanics, the vacuum is the state of minimum energy, the "ground state" full of fluctuating quantum fields.

These quantum fluctuations, ZPF a zero point energy associated with any quantum field supposedly do not involve real particles, only virtual particles-antiparticles pairs produced by the vacuum; the particle and the antiparticle try to separate for a short period (allowed by the principle of uncertainty), but they quickly annihilate with each other and the pair disappears again [36]. Since there is infinite number of harmonic oscillators per unit volume, the total zero-point energy density is, in fact, infinite. Through the process of renormalisation some sense can be made of infinite quantities. As it is usually implemented, this yields a zero energy density for the standard quantum vacuum.

The theory of Relation postulates that the zero-point energy yields a non-zero energy density for the quantum vacuum. It is extremely difficult to observe these vacuum fluctuations, since there is no state of lower energy to which the vacuum can be compared. However, the "Casimir effect" is one situation in which its effects can be seen indirectly.

In 1948, the Dutch physicist H.B.G. Casimir predicted that two clean, neutral, parallel, microscopically flat metals attract each other by a very weak force that varies inversely as the fourth power of the distance between them. It was experimentally verified in 1958. It can be understood in the following way: the zero-point energy filling the vacuum exerts pressure on everything; in most circumstances, this pressure is not noticeable, since it acts in all directions and the effect cancels. However, the quantum vacuum has different properties between the two plates. Some of the zero-point vibrations of the EM field are suppressed namely, those with wavelengths too long to fit between the plates. So, the zero-point energy density between the plates is *less* than that of the standard vacuum, *i.e.* it is negative. From this it follows that the pressure outside is greater and hence the plates feel an attractive force.

We think that it is precisely because the zero-point energy density between the plates is *less* than that of the standard vacuum that this negative energy induces (allowed by the CP) more positive energy, which is attractive. Since the beginning, the ground state of the going on expansion loses constantly some energy, and real fluctuations transform negative energy into positive energy. Pretend there is only virtual vibrations is a quantum way to corroborate the classical empty spacetime and deny physical significance to the states of negative energy.

The vacuum EM zero-point radiation is very dim, and theorists concluded that gravitation appears to be a long-range interaction akin to the van der Waals force. We prefer to declare it similar to the electromagnetism, which is less restrictive and includes both forces Coulomb and van der Waals. Roughly speaking, all the forces involved in interactions between atoms can be traced to the EM force acting on the electrically charged protons and electrons inside the atoms. This contains the intermolecular forces between the individual molecules.

On one hand, at the ZPF of the vacuum, the negative EM force of the vacuum appears to be itself a long-range interaction such as the van der Waals force because it has been diluted since the beginning: the EM charge weakens during expansion and looks like a residuum force. And, if gravity is treated as a residuum force, it is because inertial and gravitational forces alike arise due to an interaction between vacuum EM zero-point radiation and negative subatomic particles giving birth to ordinary matter [35].

On the other hand, Coulomb force, which, classically viewed, acts directly between charged particles, is enormously stronger than the gravitational force. We calculate a difference of about 10^{36} orders of magnitude between the EM and the gravitational forces between two protons in the atom like if, in microcosm, electromagnetism was the only force existing. Suppose two protons on a surface of a collapsing star due to gravity, so that at the end a little bit more than a Fermi separated them. Then the difference of strength between gravity and electromagnetism would be of the order of about 1 rather than 10^{36}. Gravitation, in this view, appears to be a short-range interaction akin to the Coulomb force. Gravity could then be construed as the second side, opposite and attractive, of the EM force. We already said that since the beginning of our universe the EM negative-energy transforms itself into positive gravitational energy. To be consistent with our theory, if we take the EM force of the vacuum and go back (from present to past) to the big bang and, on the other hand, if the whole universe collapses gravitationally (going from present to future), we will have the same magnitude of force. Like if electromagnetism in the vacuum and gravitation between stars were the repulsive and attractive sides of one force with EM origin. This said for the order of magnitude, we do not allege that gravity is exactly the opposite replica of electromagnetism.

4.2. Cosmological Constant

For a few years observational advances have led to the conclusion that the present universe is dominated by an energy density with negative pressure. Cosmologists call it dark energy and made it a version of Einstein's cosmological constant a constant, unchanging energy that pushes space to expand.

Microwave background experiments as well as supernova results strongly support the proposition coming from inflationary models that space is flat. To equal the critical density in a spatially flat universe it requires the total mass/energy density. Observations of galaxy clusters tell us that the matter contribution (baryonic and non-baryonic) to the total energy density can amount to only 20% - 30% of the needed critical density. Dark energy, which makes up about three-quarters of the energy density of the universe, accelerates cosmic expansion. With inflationary cosmology, the cosmological constant is driving this accelerated expansion only for a tiny fraction of a second. With quintessence model there is a slower and steady accelerated phase of spatial expansion. Whatever the model, a non-vanishing cosmological constant is expected, on the basis of quantum

theory, to describe the astronomical observations [37].

Cosmologists are then confronted with two problems. First, since all sorts of vacuum energies contribute to the effective cosmological constant, why is the total vacuum energy density so small by all particle physics standards? Second, since the vacuum energy density is supposed to be constant in time while the matter energy density decreases as the universe expands, what can explains this cosmic coincidence that both densities are comparable just at the present time, while their ratio has been tiny in the early universe and will become very large in the distant future? [38]

Theory of Relation is not an inflationary cosmology and claims that the energy density of vacuum is negative with an associated positive pressure driving a decelerated expansion of empty space. It assimilates dark energy with the cosmological constant, and also with the negative energy, while it is generally perceived as "exotic" positive energy. Even if the assumption was right that a spatially flat universe needs about 30% of ordinary and dark matter and about 70% of dark energy to obtain the total mass/energy density equaling the critical density, we disagree with the interpretation that the expansion of the universe is accelerated. We think the observations, supernova and CDM, are construed through a very speculative inflationary cosmology. We conclude that the percent of positive matter will grow and that the percent of expansionist dark energy will wane until equality, it will be then the path toward a big crunch.

Cosmologists are not satisfied with identifying the cosmological constant with dark energy because the expected sea level for the quantum vacuum is much higher than we observe: naively one should indeed have a characterless universe, with matter drowned by ~122 orders of magnitude below the energy density of the cosmological constant. But the cosmic concordance measures only a factor of a few differences. Furthermore, the matter and radiation we see in the universe evolves with the expansion, while the cosmological constant does not [37]. What is the explanation? The theory of Relation assumes that there is more a causal relation than a scheming coincidence encloses in the question "why both densities without apparent relation are today about the same order"? His approach is to change the initial conditions: the energy density of the cosmological constant was large at the beginning and composed of baryons and bosons which gave up their energy to become actual "tired matter" and "tired light". On this temporal mode, the primordial vacuum energy was ~122 orders of magnitude greater than the present nearly zero of the cosmological constant. The matter and light of our universe came from the transformation of the negative energy-matter into positive energy-matter throughout expansion. The energy of the vacuum is dynamic; its density varies in time: that is why the density of the global positive matter diluted into the expansionist void coincides with the density of the vacuum [39] [40].

All this negative energy is now recycled into positive matter in the framework of GR with four dimensions, with matter-galaxies attached to the fixed grid on

the expanding fabric of spacetime, like dots on the surface of a balloon. Density of matter of the whole universe is very weak, nearly zero, but would be as big as initial negative energy if the mass distributed throughout the surface was concentrated in a pre-relativistic Newtonian centre. On this spatial mode, the small, dense vacuum would produce a higher energy of 122 orders of magnitude. Subsequently, the *spatial* radius of the universe would have the same variation as the *temporal* wave of spacetime expansion: $R_u = t_o c_{em}$.

About 4% of matter is ordinary (baryonic: ~0.4% visible, ~3.6% not visible), and we speculate that dark matter is a negative matter which is still waiting to be transformed into positive ordinary matter. The very low dark energy, taking the form of a cosmological constant, is like a cold radiative surface on a featureless negative Dirac's ocean. A low vacuum energy keeps evaporating, which dulls the negative EM energy of our second structure, and in the same breath enlarges the positive energy, which allows the gravitational charge to take the ascendant. It is the deep explanation of the tired light theory.

Note that "quintessence" is dynamic, and has a density of energy that varies through time, like the negative EM spacetime wave. Quintessence is a hypothetical form of dark energy postulated as an explanation for observations that infer an accelerated universe; it is a positive energy density of the vacuum which proceeds from a positive cosmological constant with a negative pressure. By contrast, the negative EM spacetime wave of the theory of Relation has a negative energy density of vacuum resulting from a negative cosmological constant in decline, which implies a positive pressure that requires a decelerated expansion of the universe.

4.3. Principle of Compensation and the Zero-Point Fluctuations

In order to solve the "unsolved mystery in modern physics", known as inertia (the instantaneous opposition to acceleration of all material objects), Haisch and Rueda related EP to Sakharov's conjecture of a connection between Einstein action and the vacuum [41] [42]. They said that the properties of the EM quantum vacuum as experienced in a Rindler constant acceleration frame were investigated, and the existence of an energy-momentum flux was discovered and also its relative radiation, both stem from event-horizon effects in accelerating reference frames. The force of radiation pressure produced by the energy-momentum flux proves to be proportional to the acceleration of the reference frames, which leads to the *quantum vacuum inertia hypothesis* meaning that the inertia of matter could be interpreted at least in part as a reaction force originating in interactions between the EM zero-point field and the elementary charged constituents (quarks and electrons) of matter. They tried to demonstrate that this approach to inertia is consistent with GR and it appeared also that Newton's equation of motion ($f = ma$) could be inferred from Maxwell's equations as applied to the ZPF, *i.e.* the stochastic electrodynamics (SED) version of the quantum vacuum [43].

Their goal was to show that a ZPF-EM reaction force will prove to be the exact opposite of this, and can therefore reasonably be interpreted as the inertia of the object, *i.e.* that in general [41] [44]

$$f^{zp} = f_r = -f \ . \tag{19}$$

While we view their approach as promising we propose that the link of inertia of material objects with the hidden structure of the vacuum pass by both EP and CP. To illustrate this let us take a simple and gross comparison from the theory of Relation with two structures.

Imagine that the negative and expansionist EM wave of spacetime in the second hidden structure of the vacuum is represented by a rocket A going toward infinity. The two fuels of the rocket are mixed in the combustion chamber, where they undergo a violent chemical reaction. The products of this reaction are hot, high pressure gases that escape through the exhaust nozzle with a high velocity and thereby acquire a large backward momentum. The total momentum of the system must remain constant and so the backward momentum acquired by the ejected gases must be compensated by an equal and opposite forward momentum given to the rocket and unspent fuel: $p_{r(A)} = -p_{(A)}$. In terms of force: $f_{r(A)} = -f_{(A)}$. The reaction force $f_{r(A)}$ is EM and it is the asymmetric energy and momentum flux at a non-zero-point-field ZPF, so this back-reaction of lost EM energy is also a zero-point force ZPF called the inertial reaction:

$$f_{r(A)} = f_{(A)}^{zpf} \ .$$

$$-\left(-f_{(A)} = f_{r(A)} = f_{(A)}^{zpf}\right). \tag{20}$$

Imagine also that the first structure of positive gravific spacetime is represented by a similar rocket B in gravity-free space going in opposite direction; that some sort of connecting pipe (or wormhole), can collect the backward negative energy and momentum flux of A, and transform it into a positive energy which is conveyed to B. With this momentum, B will accelerate ($f = ma$) toward star and will have an equal and opposite reaction in direction of A: $-f_B = f_{r(B)}$ $\left[\left(f = ma\right); f = -f_r = -\left(-ma\right); f_r = -ma = -f\right]$. In this case we can say that the back momentum of B in direction of A is equal to the back momentum of A in direction of B: $f_{r(B)} = f_{r(A)}$, giving

$$+\left(-f_{(B)} = f_{r(B)} = f_{(B)}^{zpf}\right). \tag{21}$$

This equivalence of reactions gives apparently a connection between Einstein's EP and the inertia of the vacuum based on Sakharov's conjecture.

In the case of *stationary* orbit, there seems no difference between positive or negative energy, like if there was just one structure with no vacuum energy

$$\left[-\left(-f_{(A)} = \overrightarrow{f_{r(A)}} = f_{(A)}^{zpf}\right) = +\left(-f_{(B)} = \overleftarrow{f_{r(B)}} = f_{(B)}^{zpf}\right)\right]. \tag{22}$$

There is a balance between F_g and F_i, and the ZPF which is usually perceived as isotropic, symmetric, stationary or uniform-motion frame, without

resistance. The EP is safe but does not explain the lost energy of A gained by B.

In deep space, A and B are in a *dynamical system*. With time and distance, a kind of resistance not coming from particles of positive energy settles in front of A, which causes an unexpected loss of energy and a decrease in speed. It follows that B receives additional energy and has an abnormal acceleration. After a while, the deceleration of A became smaller, as did the acceleration of B. The EP explains neither this anomaly nor the change of sign of the energy.

We consider that Pioneer in deep space is a material object in a dynamical system which undergoes a change of its velocity with respect to time on both interpenetrated structures. The tired light wave of spacetime causes an asymmetric non-zero-point field of the quantum vacuum which decelerates the advancing probe. Tired light phenomenon is analogously the rocket A which loses energy going toward infinity.

According to the strong equivalence principle (SEP), in a zero homogeneous background of ordinary classical electrodynamics, inertial mass is equal to active and passive gravitational mass ($m_i = m_{g-a} = m_{g-p}$). Ordinary classical electrodynamics presupposes positive energy solutions and, accordingly, negative energy is merely Dirac's artifice, a virtual particles background that not changes the laws of physics when exchanging one inertial reference system for another. In a non-zero background SEP could be satisfied if

$$\overleftarrow{m_g a} = \overrightarrow{m_i a} = \overrightarrow{m_{zpf} a}. \tag{23}$$

It would be operational if at least there was equality between the coefficients of resistance m_g and m_{zpf}. This would require more dust with positive energy in deeper space which is not the case ($m_{zpf} \neq +E_{zpf}/c^2$).

On the contrary, the space vacuum density is less positive. If the ZPF force represents the system of inertia (resistance to positive energy), then there is less inertia than gravity, which is construed by the theory of Relation as a reduction in negative EM energy converted into positive gravitational energy. At point F_{ZPF}, the declining negative energy flux ($-F_{ZPF} = -E_{ZPF}/c^2$), characterized by more empty space, interacts with quarks and electrons in the physical probe, adds a growing positive energy ($+E_{zpf}/c^2$), and generates an inertia reaction force in the opposite direction. The theory includes the real, not just virtual, effects of the quantum vacuum EM in physics, and CP allows the conversion of the energies that induce weight. The anomalous deceleration of Pioneer outwards is interpreted like an anomalous acceleration toward the Sun, giving

$$f_{(P)} = -f_{r(P)} = -f_{(A)}^{zpf}. \tag{24}$$

There is an anomalous acceleration toward the Sun in relation with an equal deceleration outwards, an imbalance between $\overleftarrow{m_{g-a}}$ and $\overrightarrow{m_i}$, with $m_{g-a} > m_i$. By considering Pioneer as an integral part of a freely falling local frame in a flat space-time, an "abnormal" influence *is measurable* from the gravitational potential which causes an acceleration of the frame as a whole. The outcome of a local non-gravitational test experiment *is dependent* of the velocity and location of

the freely falling frame relative to other matter in the universe [45] [46] [47].

Relatively to Newton's third law, which requires that the driving force defined in the second law be counterbalanced by a reaction force traditionally called the inertia of matter [44], the meaning is that the abnormal acceleration of Pioneer toward the Sun is counterbalanced by the deceleration of the expansion in opposite direction.

The Pioneer's anomaly brings an argument against the equality (or proportionality) of inertial mass to gravitational mass within the standard theoretical framework of GR. The energy flux at point F_{ZPF} causes asymmetry in the metric: the gravitational *system* toward Sun gains energy ($+f_{(B)}^{zpf}$) while the inertial system outwards loses the same energy ($-f_{(A)}^{zpf}$). Equation (22) becomes an inequality

$$\left[-\left(-f_{(A)}=f_{r(A)}=f_{(A)}^{zpf}\right)-\Delta f_{(A)}^{zpf}\right]\neq\left[+\left(-f_{(B)}=f_{r(B)}=f_{(B)}^{zpf}\right)+\Delta f_{(B)}^{zpf}\right]. \quad (25)$$

It also exhibits two opposite systems, like two structures going in the opposite infinity but in close relation, one springing from the other.

In virtue of the CP there is an induced gravitation at the point ZPF: negative EM energy of the expansion converts its energy into positive gravitational condensable energy. This is reflected in a compensation which requires that the EM background has a spectrum of variable Lorentz energy density and real vacuum interactions. In that sense, both energies are equivalent to give birth to inertia.

Both principles are complementary and are necessary to understand our dual universe: EP for the global positive matter; CP for the universal negative expansion. An observer on the first structure (positive) will have the trend to adopt Mach proposition saying that inertia, the instantaneous opposition to acceleration of all objects, somehow originated in a global linkage of all matter in the universe. An observer on the second structure will adopt the view of Newton that inertia is an inherent property of matter for which a further explanation possible is a link with an absolute space. For Mach, matter is inertia. For Newton, space is inertia. Mach's explanation is for the positive energy while the one of Newton is for the negative energy. The theory of Relation embraces both propositions: absolute space of Newton is the negative EM energy of the expansion converted into the positive energy of matter.

4.4. Pioneer Effect and Theory of Quantum Gravity

If anomalies in various stellar and galactic redshifts, commonly interpreted as Doppler shifts in an expanding universe, are in fact the result of loss of energy by observed photons traversing a radiation field (Erwin Finlay-Freundlich)), where does this loss of energy go? No generally accepted physical mechanism has been proposed for this loss, which seems against the conservation of energy, and some conclude that the assumption of a universal expansion is erroneous [48].

The theory of Relation says that photons slowly lose energy across an expanding world, thus creating an EM space-time. Since a fall of energy corres-

ponds to an increase in cosmological wavelength of light, this effect would produce a redshift in the spectral lines which would augment proportionately with the distance of the source. The energy lost by the photons of the EM cosmological wave of spacetime reappears in gravific spacetime. If we state that there is an EM negative structure of the expansion directly connected with the positive gravitational structure, by which unknown physical mechanism in the vacuum the lost negative energy could be converted into positive energy?

The theory of Relation is in the line of those who support the ZPF proposal seeking to ascribe gravitation entirely to interaction with zero-point radiation. Puthoff's gravity [34] [49], suggesting that gravitation is a residual effect of the ZPF of the vacuum EM field, is positive-energy "pushing together" gravitation, similar to Lesage's "ultra-mundane corpuscles". We prefer a model proposed by Hotson [8], which also utilizes a residual effect of electromagnetism, whereas negative-energy forces act by "pulling together" [35]. This attraction suggests that the zero-point energy is felt by the Casimir effect. Two metal plates brought sufficiently close together, allowing only small, high-frequency EM "mode" of the vacuum energy to squeeze in between, will attract each other very slightly [50]. Frequencies of any type of EM mode (which give more fluctuations) outside of the plates creates a vacuum pressure which brings the two plates closer to each other. Our explanation is to say that empty space with longer wavelengths induces gravity. There is attraction when the energy is thrown out of the space which separates the plates, the energy then becomes lower than that of the standard vacuum. We assimilate this phenomenon to a kind of gravity illustrating the CP: the less energy there is in empty space, the more gravity there is.

If gravity pulls the plates together, how comes? By what mechanism if the negative energy acts? It would take a quantum theory of gravity involving the idea of the "graviton" as the quantum of the gravitational field to answer these questions. Unfortunately, it is not easy to construct theories of physics *in vacuo*. Physicists feel that attempt to formulate a theory of quantum gravity to incorporate GR and quantum mechanics in a fundamental way must require the input of radically new concepts and quantic language of gravitation. In spite of a variety of *a priori* assumptions, arguments and disagreements, not only there is no experience to test a theory of quantum gravity but there is not even a reasonable existing theory [51]. The absence of any experimental input opens up a Pandora's Box of possibilities. One approach, which remains elusive, is to regard gravity as "just another field", and to treat it in the way that has been successful with the EM interactions. Our view is that Pioneer effect could be a kind of experimental input that might provide a consistent theory of quantum gravity.

In 1937, Louis de Broglie wrote: "The neutrino would be a kind of half-photon. Isolated, that is to say not accompanied by an antineutrino, it would have no electromagnetic field since it could not annihilate itself by photoelectric effect. But united with an antineutrino, it would form a photon

and would possess an electromagnetic field of Maxwellian type" [52] [53]. Gamow [54] and Niels Bohr emitted the possibility of a link between neutrino and graviton. In the early 1950s, Erwin Finlay-Freundlich proposed a redshift as "the result of loss of energy by observed photons traversing a radiation field" [55]. R.A. Alpher noted "No generally accepted physical mechanism has been proposed for this loss", though P.F. Brown proposed that the energy lost reappears as neutrino pairs resulting from the exchange of a graviton between two photons.

For our part, we think that the empty space created by the evacuated energy between the Casimir plates is filled with gravitons which can be created by neutrinos. We suggest that "*the neutrino is a kind of half-graviton. Isolated, that is to say without an antineutrino, it would have no gravitational field since it could interact just with weak nuclear force. But linked to an antineutrino, it would form a graviton and would possess a gravitational field of Casimir type*" (The square root of a negative number provides a negative solution and, in this case, a real negative energy solution)

$$\begin{array}{llll} \text{neutrino} & \& & \text{antineutrino} \to \text{graviton} \\ 10^{-34}\,\text{kg} & \times & 10^{-34}\,\text{kg} & \to 10^{-68}\,\text{kg} \\ \left(10^{-68}\right)^{1/2}\,\text{kg} & \times & \left(10^{-68}\right)^{1/2}\,\text{kg} & \to 10^{-68}\,\text{kg} \\ 1/2\ \text{graviton} & & 1/2\ \text{graviton} & \end{array} \tag{26}$$

This assumption of a graviton composed of two half neutrinos is based on a connexion between Casimir experience, the Hubble constant, the Pioneer effect, and also with the age and the radius of the universe.

Firstly, about the Casimir experience, S. K. Lamoreaux, while at the University of Washington, conducted the most precise measurement of the Casimir effect [50]. Helped by his student Dev Sen, he used gold-coated quartz surfaces as his plates. One plate was attached to the end of a sensitive torsion pendulum; if that plate moved toward the other, the pendulum would twist. A laser could measure the twisting of the pendulum down to 0.01 micron accuracy. A current applied to a stack of piezoelectric components moved one Casimir plate; an electronic feedback system countered that movement, keeping the pendulum still. Zero-point energy effects showed up as changes in the amount of current needed to maintain the pendulum's position. Lamoreaux found that the plates generated about 100 microdynes (one nanonewton: 10^{-9} N) of force. The result falls within 5% of Casimir's prediction for that particular plate separation and geometry. The strength generated by the plates is

$$hc/\lambda^2 = 10^{-9}\,\text{N}, \tag{27}$$

so $\lambda = 1.41\times10^{-8}\,\text{m}$. Energy is $F\lambda = \left(10^{-9}\,\text{N}\right)\left(1.41\times10^{-8}\,\text{m}\right) = \sim 1.4\times10^{-17}\,\text{J} = \sim 88\,\text{eV}$. The force of the Casimir experience moving a plate toward the other comes into existence from the energy of quantum fluctuations, from physical properties of the vacuum [56].

Secondly, if we suppose that the configuration of all the mass in the universe,

estimated ~2×10^{53} kg, is concentrated at the center of a sphere with radius ~10^{26} m, we note about the same Casimir strength. The force between the global mass and a peripheral body of 1 kg would be ~10^{-9} N:

$$F = GMm/R^2 = \left(2 \times 10^{53} G \times 1\right) \Big/ \left(10^{26}\right)^2 . \quad (28)$$

The gravitational acceleration toward the center would be ~10^{-9} m/s^2.

Thirdly, the cosmological acceleration ~8×10^{-10} m·s^2, noted before (18), is also the reported Sunwards anomalous acceleration of Pioneer 10 [5]. There is a correlation between 1) the acceleration of expansion with Hubble constant based on the age of the universe ($H = 1/t$), 2) the anomalous acceleration of the spacecraft, and 3) the acceleration of time [57]. On one hand, 2) and 3) have the same acceleration and force in common, and are roughly on the same "cosmological" scale. On the other hand, even if the three items have the same force (10^{-9} N), the microscopic scale of Casimir force and the macroscopic scale of the universe have apparently no rational classical link. With the theory of Relation there is a deep and strong link. Both effects, Pioneer and Casimir, happen in empty space, at the zero-point energy. Concerning Pioneer effect, we have already put a link between the negative EM cosmological wave of spacetime losing energy and the gravity which induces matter. In the case of Casimir, the effect happens after a void is created between the plates. We assume that the expulsion of the positive energy which creates a void gives way to the negative scalar EM cosmological wave of spacetime which is on the same scale as 2) and 3) (*i.e.* the EM cosmological wave equivalent to the radius of our universe.). Since there is almost the same strength between the anomalous acceleration of the Pioneer effect and the Casimir effect in empty space, we presume that the void in Casimir experience is about the same magnitude as that of the vacuum enveloping the spacecraft. At this zero-point energy, there is an ocean of photon and electrons and, also, of gravitons and neutrinos.

The force ~10^{-9} N corresponds to a particle ~10^{-34} kg (~88eV) considered as a possible mass for the neutrinos. Our proposition for the neutrino to be half a graviton gives for the graviton

$$\left(\sim 10^{-34}\,\text{kg}\right)^2 = \sim 10^{-68}\,\text{kg}; \left(\sim 10^{-33}\,\text{eV}\right). \quad (29)$$

This weight matches with the one assessed for the graviton particle [58], if we assimilate the negative EM cosmological wave with the radius of the universe (based on Schwarzschild: $R_{SCH} = t_o c = GM/c^2$); $t_o c$ is then like a Compton wavelength of the quantum graviton for the radius of the universe

$$\lambda_g = h/m_g c = c/H = t_o c = R_U = \sim 10^{26}\,\text{m}. \quad (30)$$

The value ~10^{-68} kg for m_g could also be for a Compton wavelength of the quantum photon for the radius of the universe

$$\lambda_\gamma = h/m_\gamma c = c/H = t_o c_{EM} = R_U = \sim 10^{26}\,\text{m}. \quad (31)$$

Note that $m = \sim 10^{-68}\,\text{kg}\left(\sim 10^{-33}\,\text{eV}\right)$ could be a quantum graviton, or photon, and the dimensionless coupling constant for the universe could be $GM^2_{\text{graviton(or photon)}}/\hbar c = \sim 10^{-122}$. The inverse is the cosmological constant: $=\sim 10^{122}$ [59].

Pioneer effect, tired light, and the postulated graviton resulting from neutrinos pair in the framework of the theory of Relation, could provide a basic idea of what we simply do not know.

5. Conclusions

General Relativity deals with the interaction between a system in inertial space and gravity-matter based on the principle of Equivalence. The theory of Relation considers it like its first structure of "condensation" with a positive matter-energy accompanied by an empty space generally ascribed without gravitation or energy. This structure is directly counterbalanced by a second one with a negative matter-energy driving the expansion. Since the so-called big bang, a negative electromagnetic mass-energy transforms itself into the positive gravitational mass-energy. We live in a *dualverse* rather than a universe, where the parameters of both structures are not independent but linked by a communicating vessels system, such as all that is earned by one is lost by the other one. One of the consequences is a deceleration of the expansion in phase with a local increment of gravity of the galaxies. Since 1998, a logic and over-simple line of reasoning based on a positive dark energy of empty space, interprets wrongly, in our opinion, the global expansion as accelerated with a positive cosmological constant.

The Pioneer anomaly in deep space is the first to reveal a decreasing negative electromagnetic wave of tired light which induces a positive gravity in the vacuum, by virtue of the postulated principle of Compensation. We think that other anomalies, like "Allais effect" [60] [61] or "galaxy rotation curves" [62], can also be explained by the theory of Relation, which is in the frame of "cyclic universe theory".

In the deep space outside our solar system, the probe is in a flat space (not in a Hubble space with recession), in the furrow of an expansible electromagnetic wavelength of spacetime. With the distance in vacuum, the minimal negative electromagnetic energy slowly lowers which induces the positive gravitational energy. In this induction process, cosmological photons gradually lose energy, thus heightening the gravitational potential. It is like if the electromagnetic energy was converted into gravitational resistance or acceleration toward matter. Pioneer 10 slows down going outwards and the observed frequency of the redshift on Earth is lower than predicted. Pioneer effect is an experimental "violation" of the principle of Equivalence. Such a violation shakes the foundation of physics. That does not mean that Einstein's theory is fundamentally "wrong" but incomplete, that the gravitational interaction is more complex than previously assumed, and contains, in addition to the interacting Einsteinian field of spin 2,

the effect of another long-range field, which is that of the second structure which incorporates an electromagnetic spacetime. This new vision of the universe is based on the Compensation principle allowing interaction between material system in deep space and vacuum electromagnetic-zero-point fluctuations inducing gravity. The equivalence principle and the compensation principle coexist and are complementary.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Anderson, J.D., Laing, P.A., Lau, E.L., Liu, A.S., Nieto, M.M. and Turyshev, S.G. (1998) *Physical Review Letters*, **81**, 2858-2861. https://doi.org/10.1103/PhysRevLett.81.2858

[2] Turyshev, S.G., Anderson, J.D., Laing, P.A., Lau, E.L., Liu, A.S. and Nieto, M.M. (1999) The Apparent Anomalous, Weak, Long-Range Acceleration of Pioneer 10 and 11. 34*th Rencontres de Moriond Meeting on Gravitational Waves and Experimental Gravity*, Les Arcs, 23-30 January 1999, 1-7. https://doi.org/10.2172/353450

[3] Anderson, J.D., Lau, E.L., Turyshev, S.G., Laing, P.A. and Nieto, M.M. (2002) *Modern Physics Letters A*, **17**, 875-885.

[4] Mbelek, J.P. (2003) The Pioneer Anomaly: A Bulk Scalar Field? Service d'Astrophysique, C.E. Saclay F-91191 Gif-sur-Yvette Cedex, 13.

[5] Anderson, J.D., Laing, P.A., Lau, E.L., Liu, A.S., Nieto, M.M. and Turyshev, S.G. (2002) *Physical Review D*, **65**, Article ID: 082004.

[6] Markwardt, C.B. (2002) Independent Confirmation of the Pioneer 10 Anomalous Acceleration.

[7] Dittus, H., Turyshev, S.G., Lämmerzahl, C., Theil, S., Foerstner, R., Johann, U., Ertmer, W., *et al.* (2005) A Mission to Explore the Pioneer Anomaly. *Proceedings of* 39*th ESLAB Symposium*, Noordwijk, 19-21 April 2005, 1-9.

[8] Hotson, D.L. (2002) Dirac's Equation and the Sea of Negative Energy. *Infinite Energy*, Issue 44, Part 1, 10-12, Part 11, 2, 6, 19-22.

[9] Hatch, R.R. (1999) *Electrodynamics*, **10**, 69-75.

[10] Einstein, A. (1961) Relativity. Three Rivers Press, New York.

[11] Ostermann, P. (2003) Relativity Theory and a Real Pioneer Effect.

[12] Born, M. (1962) Einstein's Theory of Relativity. Dover Publications, Inc., Mineola, 121-125, 216, 339, 352-353.

[13] Rocard, J.-M. (1986) Newton et la Relativité, Que sais-je? Presse universitaire de France, 103.

[14] Stolmar, A. (2001) Pioneer 10 Anomalous Acceleration Is Distance Caused Hubble Red-Shift. http://sci.tech-archive.net/Archive/sci.physics/2006-01/msg00766.html

[15] Atkins, K.R. (1970) Physics. J. Wiley & Sons, Inc., Hoboken, 130.

[16] Crawford, D.F. (1999) A Possible Explanation for the Anomalous Acceleration of Pioneer 10.

[17] Didon, N., Perchoux, J. and Courtens, E. (1999) Université de Montpellier.

[18] Nieto, M.M., Turyshev, S.G. and Anderson, J.D. (2005) *Physics Letters B*, **613**, 11. https://doi.org/10.1016/j.physletb.2005.03.035

[19] Tcherepanov, O.A. (2002) Doppler-Mikhelson's Principle and a Pseudacceleration of the Nasa's Spacecrafts: Pioneer 10 and Pioneer 11 Have Discovered in the Perihelion Space the Faintly Refractive Medium.

[20] Ratcliffe, H. (2005) *Progress in Physics*, **3**, 19-24.

[21] Thuan, T.X. (1995) Un Astrophysicien. Champs Flammarion, Paris, 76-83.

[22] Dive, P. (1945) Les Interprétations Physiques de la Théorie d'Einstein. Dunod, Paris, 52.

[23] Renshaw, C.E. and Kallfelz, W.L. (1998) *Google Scholar*, *Physics Essays*, **11**, 361-366.

[24] Einstein, A. (1972) Réflexions sur l'électrodynamique, l'éther, la géométrie et la relativité. Gauthier-Villars Éditeur, 73, 113-114.

[25] Johann, R., Dittus, H. and Lämmerzahl, C. (2007) Exploring the Pioneer Anomaly: Concept Considerations for a Deep-Space Gravity Probe Based on Laser-Controlled Free-Flying Reference Masses.

[26] Zwicky, F. (1929) *PNAS*, **15**, 773-779. https://doi.org/10.1073/pnas.15.10.773

[27] Hubble, E. and Tolman, R.C. (1935) *Astrophysical Journal*, **82**, 302. https://doi.org/10.1086/143682

[28] Pecker, J.-C. and Vigier, J.-P. (1986) A Possible Tired Light Mechanism. *Observational Cosmology*, *Proceedings of the IAU Symposium*, Beijing, 25-30 August 1986, 507 p. https://doi.org/10.1007/978-94-009-3853-3_48

[29] Tolman, R.C. (1934) Relativity, Thermodynamics and Cosmology. Dover Books, Mineola, 287.

[30] Misner, C., Thorne, K. and Wheeler, J.A. (1971) Gravitation. Freeman, New York, 426-428.

[31] Sakharov, A.D. (1967) *Doklady Akademii Nauk SSSR*, **177**, 70-71. (English Translation (1968) in *Soviet Physics—Doklady*, **12**, 1040-1041).

[32] Zeldovich, Yu.B. (1967) *Zh. Eksp. & Teor. Fiz. Pis'ma*, **6**, 883-884. (English Translation (1967) in *Soviet Physics—JETP Letters*, **6**, 316-317)

[33] Rueda, A., Haisch, B. and Tung, R. (2001) Gravity and the Quantum Vacuum Inertia Hypothesis. 1. Formalized Groundwork for Extension to Gravity.

[34] Puthoff, H.E. (1989) Gravity as a Zero-Point-Fluctuation Force. https://doi.org/10.1103/PhysRevA.39.2333

[35] Ridgely, C.T. (2001) Gravitation and Forces Induced by Zero-Point Phenomena.

[36] Coughlan, G.D. and Dodd, J.E. (1991) The Quantum Vacuum. In: *The Ideas of Particle Physics*, Cambridge University Press, Cambridge, 33-34.

[37] Greene, B. (2005) The Fabric of the Cosmos: Space, Time and the Texture of Reality. Vintage, New York, 434-435.

[38] Straumann, N. (2002) On the Cosmological Constant Problem and the Astronomical Evidence for a Homogeneous Energy Density with Negative Pressure.

[39] Jodra, S. (2002) La face obscure de l'expansion cosmique. *Science & Vie*, *Hors-Série*, No. 221, 130.

[40] Lachièze-Rey, M. (2004) La constante cosmologique. *Pour la Science*, *Dossier*, Dec., 98.

[41] Rueda, A. and Haisch, B. (1998) *Physics Letters A*, **28**, 1057. https://doi.org/10.1023/A:1018893903079

[42] Rueda, A. and Haisch, B. (1998) *Physics Letters A*, **240**, 115-126. https://doi.org/10.1016/S0375-9601(98)00153-4

[43] Sunahata, H. (2005) Interaction of the Quantum Vacuum with an Accelerated Object and Its Contribution to Inertia Reaction Force. https://www.researchgate.net/publication/241528772

[44] Rueda, A. and Haisch, B. (1998) *Foundations of Physics*, **28**, 9, 19.

[45] Will, C. (1989-1998) The Renaissance of General Relativity. In: *The New Physics*, Cambridge University Press, Cambridge, 7-33.

[46] Barry, P. (2007) Testing the Equivalence Principle ["Mr. Galileo Was Right."], Physorg, May 21. Quotation of C. Will Saying That EP Could Fail at Subtle Level.

[47] Jaekel, M.T. and Reynaud, S. (2005) Gravity Tests in the Solar System and the Pioneer Anomaly.

[48] Gouin, R.Y. (2004) An Astrophysical Origin to Anomalous Spacecraft Accelerations. https://www.academia.edu/24130570

[49] Haisch, B. and Rueda, A. (1997) *The Astrophysical Journal*, **488**, 563-565. https://doi.org/10.1086/304731

[50] Yam, P. (1997) *American Scientific*, **277**, 83. https://doi.org/10.1038/scientificamerican1297-82

[51] Feynman, R. (1987) Lumière et matière. InterÉditions Points-Sciences, Paris, 195-196.

[52] de Broglie, L. (1937) Matière et Lumière. Edition Albin Michel, Paris, 109.

[53] Cluny, H. (1959) La Science Nucléaire. Buchet/Chastel, Paris, 84.

[54] Gamow, G. (1962) Gravitation. P-B Payot, Paris, 142-143.

[55] Alpher, R.A. (1962) *Nature*, **196**, 367-368. https://doi.org/10.1038/196367b0

[56] Reynaud, S. (1997) L'énergie du Vide. *Science & Avenir*, No. 112, 45.

[57] Ranada, A.F. (2004) *Foundations of Physics*, **34**, 1955-1971.

[58] Valev, D. (2008) *Journal Aerospace Research in Bulgaria*, **22**, 68-82. http://arxiv.org/abs/hep-ph/0507255

[59] Demaret, J. (1991) Univers. Le Mail, 138.

[60] Allais, M. (1997) L'Anisotropie de l'Espace. Edition Clément Juglar.

[61] Bagdoo, R. (2009) Link between Allais Effect and General Relativity's Residual Arc during Solar Eclipse. http://vixra.org/abs/1302.0089 https://www.academia.edu/5537473

[62] Bagdoo, R. (2012) Galaxy Rotation Curves Traced Out by the Theory of Relation. http://vixra.org/abs/1304.0079 https://www.academia.edu/5539791

Journal of Modern Physics, 2020, 11, 648-667
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Density Profiles of Gases and Fluids in Gravitational Potentials from a Generalization of Hydrostatic Equilibrium

Richard B. Holmes

The Boeing Company, Chicago, USA
Email: rholmes001@aol.com

How to cite this paper: Holmes, R.B. (2020) Density Profiles of Gases and Fluids in Gravitational Potentials from a Generalization of Hydrostatic Equilibrium. *Journal of Modern Physics*, **11**, 648-667.
https://doi.org/10.4236/jmp.2020.115042

Received: March 13, 2020
Accepted: May 5, 2020
Published: May 8, 2020

Abstract

It is well-known that equilibrium density profiles of gases and fluids in gravitational potentials have an r^{-1} dependence, where r is the radius. This is found in both astronomical observations and detailed simulations in spherically-symmetric geometries. It is also well-known that the standard equation for hydrostatic equilibrium does not produce such solutions. This paper utilizes a Lagrangian formulation that produces a closed-form r^{-1} solution and identifies the needed terms to be added to the standard equation for hydrostatic equilibrium to obtain such a solution. Variants of the r^{-1} solution avoid a density singularity at the origin and a total-enclosed mass singularity at infinity. The resulting solutions are shown to be in good agreement with well-established density profiles of ordinary matter in galaxies, dark matter haloes, and the atmosphere of earth. Comparisons are made to solutions based on the standard hydrostatic equation, including solutions based on the Lane-Emden equation. The origins of differences are explained.

Keywords

Density Profiles, Gravitational Potentials, Dark Matter

1. Introduction

Early numerical solutions for dark matter (DM) density profiles were found to have dependence inversely proportional to radius [1]. Astronomical observations of both ordinary matter and dark matter have also inferred density profiles closely resembling the r^{-1} profile, where r is the radius from the nominal center of the distribution, in a region surrounding the origin [2] [3] [4] [5] [6]. However, the profile has several serious drawbacks. First, it gives infinite density at the

origin. Second, it gives infinite total enclosed mass. Third, it is not consistent with many observations of the shape of the inner profiles of low-spectral brightness galaxies [7] [8] [9]. Nonetheless, the r^{-1} profile is "ubiquitous" [10]. This paper will provide theoretical support that such a profile is indeed ubiquitous when there is thermal and mechanical equilibrium.

To analyze the spatial distribution of a gas of particles, N-body simulations [1] [11] or solutions to the Vlasov equation [12] are typically used. Two simplified governing equations are considered for DM density profiles in this paper. The first is the standard equation for hydrostatic equilibrium in a spherically-symmetric geometry:

$$(1/\rho)(\mathrm{d}P/\mathrm{d}r) = -m_v M_{enc}(r) G/r^2 . \quad (1)$$

here ρ is the number density of DM, P is the pressure, r is the radius, m_v is the mass of a DM particle, $M_{enc}(r)$ is the enclosed mass, and G is the gravitational constant. This equation can be solved using the well-known Lane-Emden formulation [13], which uses the assumption that pressure is a function of density only:

$$P = c_\gamma \rho^\gamma , \quad (2)$$

where c_γ is a constant for a given polytropic exponent γ. An inhomogeneous form of the Lane-Emden equation can be used when ordinary matter (OM) is present as shown later in the paper. As shown in Section 4, calculations using Equations (1) and (2) are inconsistent with an r^{-1} solution. Such solutions have no cusp at the origin and are a poor match to the standard Sersic or Einasto profiles based on observations [1] [2] [3] [4] [6] [14] [15] [16]. This is found to be true for any polytropic exponent between about 1.1 and 2 in Section 4 below.

To address this, a generalization of Equation (1) is derived using a Lagrangian formulation. This derivation is given in Section 2. A partial validation of this derivation is given in Section 3 using the properties of a well-known density profile of a gas in a gravitational potential, namely, the earth's atmosphere. Section 4 presents behavior of solutions near the origin. Section 5 presents the behavior of solutions far from the origin and presents a sample calculation for a dark-matter halo. Section 6 compares the solutions to Einasto and de-projected Sersic profiles. Section 7 compares the solutions to results from the Lane Emden equation. Section 8 provides a brief summary.

2. Derivation of a Generalized Equation for Hydrostatic Equilibrium

The assumed kinetic and potential energy densities for a medium in a spherically symmetric geometry under the influence of gravity is given by

$$K.E.(r) = P(r),\ P.E.(r) = -\rho(r) m_v M_{enc}(r) G/r , \quad (3)$$

where the variables are defined as in Equation (1). In the topic of this paper, $M_{enc}(r)$ consists of both ordinary and dark matter. The ordinary (radiant)

matter has mass density denoted $\rho_{m,rad}(r)$. Note that the pressure is treated as a kinetic energy for the obvious reason. This will allow a variational approach with a Lagrangian density. Explicit relativistic effects and the possible impact of angular momentum are neglected in this initial analysis

The pressure is assumed to be a function of density only. As an example, the number density for fermions in Equation (3) assuming thermodynamic equilibrium is

$$\rho = n_s \Big/ \left(2\pi^2\hbar^3\right) \int p^2 \mathrm{d}p \Big/ \left[\exp\left(\left\{ \left[(pc)^2 + \left(m_v c^2\right)^2 \right]^{1/2} - m_v c^2 - \mu_v \right\} \Big/ kT \right) + 1 \right], \quad (4)$$

where p is the fermion momentum, m_v is its mass, n_s is the number of spin states, μ_v is the chemical potential, and $\hbar$ is Planck's constant. The other variables have been defined earlier. In Equation (4), the chemical potential is set to zero for non-interacting particles. With Equation (4) the pressure for fermions has the usual form in the non-relativistic limit, and is given by

$$P = c_{5/3}\rho^{5/3}, \quad (5)$$

where $c_{5/3} = 1.914\hbar^2/m_v$ assuming 2 spin states for a fermion. For the relativistic case, the exponent is 4/3, and $c_{4/3} = 5.536\hbar c$. For the general case of arbitrary polytropic exponent γ and assuming c_γ is of the form $\alpha\hbar^a c^b m_v^c$ with α a dimensionless constant of order unity, one obtains $c_\gamma = \alpha\hbar^{3(\gamma-1)} c^{5-3\gamma} m_v^{4-3\gamma}$. With this relation and $m_v = 1\,\mathrm{eV}/c^2$, for example, one obtains $c_{5/3} = 7.42\times10^{-32}\,\mathrm{J\cdot m^{-2}}$ For $\gamma = 1$ and 2 one finds $c_1 \approx m_v c^2 = 16\times10^{-20}\,\mathrm{J}$ and $c_2 \approx \hbar^3/\left(m_v^2 c\right) = 1.22\times10^{-39}\,\mathrm{J\cdot m^3}$, respectively.

This initial general development also considers independent variations of pressure with density and temperature, $(\partial P/\partial\rho)_T$ and $(\partial P/\partial T)_\rho$ Moving on to the second term of Equation (3), the gravitational potential energy term is of the usual form, involving the enclosed mass computed starting from the center of mass of the overall mass distribution. It also has units of energy density. One may use the kinetic and potential energy density expressions of Equation (3) to find the density of DM or OM versus radius in a spherically-symmetric geometry for a single species of matter. This is done by taking variations of the Lagrangian density, *K.E.-P.E.* with respect to ρ and T. The result for DM or OM is (with ordinary matter or dark matter held fixed, respectively)

$$\begin{aligned} &\left[(\partial P/\partial\rho)_T + m_v M_{enc}(r) G/r \right] \delta\rho(r) + 4\pi\rho m_v^2 \left[\int_0^r r'^2 \mathrm{d}r' \delta\rho(r') \right] G/r \\ &+ (\partial P/\partial T)_\rho \, \delta T(r) = 0. \end{aligned} \quad (6)$$

This equation should hold for arbitrary small variations $\delta\rho(r)$ and $\delta T(r)$ at each point r in mechanical and thermodynamic equilibrium at temperature $T(r)$ and density $\rho(r)$. The two last terms look like they could pose a problem. To address the last term, assume Equation (2) applies, so that T is a function of ρ only, and $\delta T(r) = (\mathrm{d}T/\mathrm{d}\rho)\delta\rho(r)$. Then

$$\begin{aligned}&\left[\left(\partial P/\partial\rho\right)_T+\left(\partial P/\partial T\right)_\rho \mathrm{d}T/\mathrm{d}\rho+m_v M_{enc}\left(r\right)G/r\right]\delta\rho\left(r\right)\\&+4\pi\rho m_v^2\left[\int_0^r r'^2\mathrm{d}r'\delta\rho\left(r'\right)\right]G/r.\end{aligned}\tag{7}$$

Note that the first two terms are $\mathrm{d}P/\mathrm{d}\rho$. To address the last term, differentiate Equation (7) with respect to r to obtain

$$\begin{aligned}&\left[\mathrm{d}/\mathrm{d}r\left(\partial P/\partial\rho\right)+4\pi r^2 m_v\left(\rho m_v+\rho_{m,rad}\right)G/r-m_v M_{enc}\left(r\right)G/r^2\right.\\&\left.+4\pi\rho m_v^2 r^2\,G/r\right]\delta\rho\left(r\right)+\left[\mathrm{d}P/\mathrm{d}\rho+m_v M_{enc}\left(r\right)G/r\right]\mathrm{d}\delta\rho\left(r\right)/\mathrm{d}r\\&+\left[\mathrm{d}/\mathrm{d}r\left(\rho/r\right)\right]m_v^2 G\left[4\pi\int_0^r r'^2\mathrm{d}r'\delta\rho\left(r'\right)\right].\end{aligned}\tag{8}$$

The integral in the last term in Equation (8) can then be eliminated using Equation (7) to give

$$\begin{aligned}&\left[\mathrm{d}/\mathrm{d}r\left(\partial P/\partial\rho\right)+4\pi r^2 m_v\left(2\rho m_v+\rho_{m,rad}\right)G/r-m_v M_{enc}\left(r\right)G/r^2\right]\delta\rho\left(r\right)\\&+4\pi\rho m_v^2 r^2\,G/r\Big]\delta\rho\left(r\right)+\left[\mathrm{d}P/\mathrm{d}\rho+m_v M_{enc}\left(r\right)G/r\right]\mathrm{d}/\mathrm{d}r\left(\delta\rho\left(r\right)\right)\\&-\left(\rho/r\right)^{-1}\left[\mathrm{d}/\mathrm{d}r\left(\rho/r\right)\right]\left[\mathrm{d}P/\mathrm{d}\rho+m_v M_{enc}\left(r\right)G/r\right]\delta\rho\left(r\right)=0.\end{aligned}\tag{9}$$

The second term in Equation (9) is the change in the potential energy of the system at radius r due to mass at radius r, and can be dropped if the enclosed mass is understood to be the enclosed mass strictly less than r.

To address the term involving $\mathrm{d}/\mathrm{d}r\left[\delta\rho\left(r\right)\right]$ the second line in Equation (9) can be integrated by parts, accounting for the volume integral of the Lagrangian density, as well as the r^2 factor that is associated with it. After considerable algebra, the result is

$$\begin{aligned}&\left\{\left[\mathrm{d}P/\mathrm{d}\rho+m_v M_{enc}\left(r\right)G/r\right]\left(1/\rho\right)\mathrm{d}\rho/\mathrm{d}r\right.\\&\left.+\left[m_v M_{enc}\left(r\right)G/r+\mathrm{d}P/\mathrm{d}\rho\right]/r\right\}\delta\rho\left(r\right)=0.\end{aligned}\tag{10}$$

This equation now has the required form so that the expression in curly brackets is zero:

$$\begin{aligned}&\left(1/\rho\right)\mathrm{d}P/\mathrm{d}r+\left(1/\rho\right)\left(\mathrm{d}\rho/\mathrm{d}r\right)\left[m_v M_{enc}\left(r\right)G/r\right]\\&=-m_v M_{enc}\left(r\right)G/r^2-\left(\mathrm{d}P/\mathrm{d}\rho\right)/r.\end{aligned}\tag{11}$$

Equation (11) is the result for strict equilibrium for a medium with a pressure that depends on density only in a spherically-symmetric geometry, and with a single species of matter. The limiting case of the standard hydrostatic equation is obtained by neglect of the second terms on both sides of Equation (11).

Based on the usual derivation of Equation (1), the standard hydrostatic equation should be valid when mechanical equilibrium applies but the stricter action equilibrium of Equation (11) does not apply. The term involving $\left(1/\rho\right)\left(\mathrm{d}\rho/\mathrm{d}r\right)$ in Equation (11) arises because of the energy required to maintain a density gradient in a gravity potential, and the term involving $-\left(\mathrm{d}P/\mathrm{d}\rho\right)/r$ arises because of the form of the divergence of the radial part of the bulk modulus in a spherical geometry.

It is also worth noting that Equation (11) can be written as

$$\mathrm{d}\ln(\rho)/\mathrm{d}\ln(r) = -1. \tag{12}$$

This simplification indicates a solution for density of the form

$$\rho \sim \rho_0 (r_0/r). \tag{13}$$

This result is both remarkable and perhaps expected. It arises because of the spherical geometry and the r^{-1} potential as discussed above. It matches the Navarro-Frenk-White (NFW) result [1] as well as other related results near the origin [2] [3] [5]. However, it has two serious drawbacks. First, it gives infinite density at the origin. Second, it gives infinite total enclosed mass. One may attempt to address these issues within the Lagrangian formalism by introducing a density constraint using a Lagrange multiplier λ_ρ, and a total particle number (mass) constraint with multiplier λ_M. With these two constraints, Equation (10) becomes

$$\begin{aligned}
&\left\{\left[1+r\lambda_M/\left(\rho m_v^2 G\right)\right]^{-1}\left[(1/\rho)\mathrm{d}\rho/\mathrm{d}r\right]\left[\mathrm{d}P/\mathrm{d}\rho + m_v M_{enc}(r) G/r + \lambda_\rho\right]\right\}\delta\rho(r)\\
&+\left(\left\{2-\left[1+r\lambda_M/\rho m_v^2 G\right]^{-1}\right\}\left[\mathrm{d}P/\mathrm{d}\rho + m_v M_{enc}(r) G/r + \lambda_\rho\right]/r\right)\delta\rho(r)\\
&+\left[\lambda_M\left(4\pi m_v^2 G\right)\mathrm{d}/\mathrm{d}r(\rho/r)\right]\int_r^\infty \mathrm{d}r' r'^2 \delta\rho(r')/\left[\lambda_M + \rho m_v^2 G/r\right] = 0.
\end{aligned} \tag{14}$$

The last term arises from the variation of the total-mass constraint with respect to density. For r small (close to the origin) or large (far from the origin), one might expect this integral to be approximately zero, since the total mass constraint requires that the radial integral from zero to infinity of the density variations be identically zero. This version of the equation will be explored in subsequent sections of this paper.

3. Generalized Equation for Earth's Atmospheric Density Profile with Altitude

As a basic check of the results of Section 2, one may consider the case of earth's atmosphere. In the earth's atmosphere, it is well-known that (a) the temperature is not constant with altitude in the troposphere [17], and (b), that despite (a), a density profile computed with an isothermal atmosphere is a decent approximation. This apparent inconsistency is investigated in this section with various solutions and in the process provides some confirmation that the generalized hydrostatic equation has validity.

First the version of the generalized equation for hydrostatic equilibrium given by Equation (11) is considered. For the earth's atmosphere at distance z above sea level, one has

$$(1/\rho)\mathrm{d}P/\mathrm{d}r + (1/\rho)(\mathrm{d}\rho/\mathrm{d}z) m_{air} g z = -m_{air} g - \mathrm{d}P/\mathrm{d}\rho/(r_e + z). \tag{15}$$

here m_{air} is the atomic weight of air, approximately (29.0) (1.67×10^{-27}) kg, and g is the gravitational constant at or near earth's surface, 9.8 m·sec^{-2}. The radius r_e is measured from earth center. Next consider an equation of state

$$P(z)=k\rho(z)T(z), \tag{16}$$

where $T(z)$ is the temperature, which is treated as an input from [17] for some of the following equations. Note that $(\mathrm{d}P/\mathrm{d}\rho)/r_e = kT(z)/r_e$ is about 6.3 × 10^{-28} J·m^{-1} and that $m_{air}g = 4.7\times10^{-25}\,\mathrm{J}\cdot\mathrm{m}^{-1}$. Hence the last term of Equation (15) can be neglected relative to $m_{air}g$. Using expression (16) and Equation (15) one obtains

$$(k/\rho)\left[\mathrm{d}\rho/\mathrm{d}zT(z)+\rho(z)\mathrm{d}T/\mathrm{d}z\right]+(1/\rho)(\mathrm{d}\rho/\mathrm{d}z)m_{air}gz=-m_{air}g\,. \tag{17}$$

This equation simplifies to (neglecting terms of order z/r_e),

$$(1/\rho)(\mathrm{d}\rho/\mathrm{d}z)=\left[-m_{air}g-k\,\mathrm{d}T/\mathrm{d}z\right]/\left[kT(z)+m_{air}gz\right]. \tag{18}$$

This can be solved analytically:

$$\rho(z)=\rho_0 kT(z=0)/\left[kT(z)+m_{air}gz\right]. \tag{19}$$

here ρ_0 is the density at sea level, which is an input. The solution of Equation (19) gives poor agreement with measured data, with significantly larger density than measured at higher altitudes. This will be seen to arise from the underlying assumption of strict action (mechanical and thermal) equilibrium. Equation (19) may be compared to the isothermal approximation, which amounts to retention of only the first terms on both sides of Equation (17) and assuming constant temperature. One obtains the usual exponential atmosphere in this case,

$$\rho(z)=\rho_0\exp\left[-m_{air}gz/kT_0\right], \tag{20}$$

where T_0 is the temperature of the atmosphere at sea level.

There is also the adiabatic equation for pressure and density. The above Equations (17)-(19) assumes strict action equilibrium. The adiabatic equation assumes that hot parcels of air near the surface of the earth rise without significant heat exchange with the surrounding air. Hence it corresponds to a process in quasi-static equilibrium that is not in strict mechanical and thermal equilibrium at all layers. With this insight, the adiabatic result can be derived from Equation (17) by neglect of the last terms on both sides of the equation, and assuming the gas satisfies a polytropic relation (2) with a polytropic exponent $\gamma\sim1.4$. With these assumptions, one obtains the result from the standard hydrostatic equation, which is the well-known formula for density [18]:

$$\rho(z)=\rho_0\left[1-m_{air}gz/\left(c_pT_0\right)\right]^{1/(\gamma-1)}. \tag{21}$$

here c_p is the specific heat at constant pressure of the atmosphere at sea level. Note that $c_pT_0=\left[(\gamma-1)/\gamma\right]P_0/\rho_0$, where P_0 is the pressure at sea level. For **Figure 1** below, the value of c_p used is 1004.7 J·kg^{-1}·K^{-1}, the value of T_0 used is 290.1 K. The latter is consistent with the ideal gas law and the sea-level values of pressure and density given in [17]. For **Figure 1**, a polytropic exponent of 1.38 rather than 1.40 is used and is justifiable because of the presence of significant water vapor in the middle atmosphere between 1 and 15 km. The integration step size is 250 m.

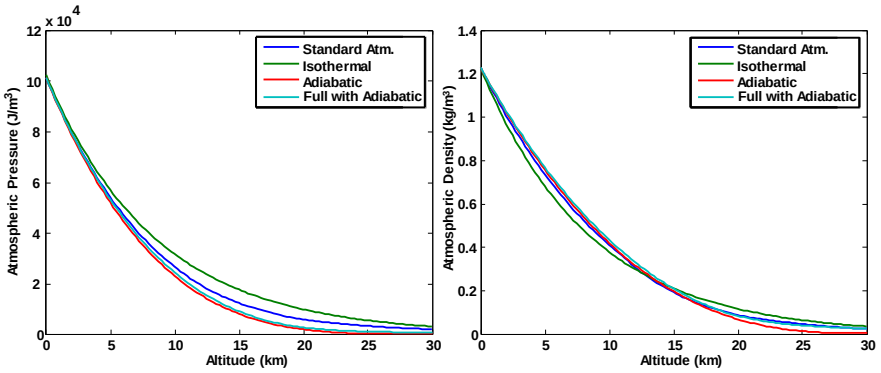


Figure 1. Solutions for Earth's atmospheric pressure (a) and density (b) versus altitude under various assumptions. Blue: standard atmospheric density (measured). Green: isothermal profile from Equation (20) with $T = 290.1$ K. Red: Adiabatic result (21) using the standard hydrostatic equation. Cyan: Adiabatic result using generalized hydrostatic Equation (15) with factor given by Equation (22).

It should be noted that Equation (21) gives nonsensical results for $m_{air}gz/(c_pT_0) > 1$, which corresponds to an altitude z of about 30 km. However, the result for density is quite accurate for the lowest 15 km of the atmosphere. To remedy the errors above 15 km, the additional terms in Equation (15) can be restored gradually above 15 km, representing the return to true equilibrium at higher altitudes. The altitude dependent factor used to restore the full Equation (15) is given by Equation (22):

$$f(z) = H(z-15)\left\{1-\exp\left[-(z-15)^2/20^2\right]\right\}, \tag{22}$$

where z is in km in this expression and $H(z)$ is the Heaviside function. This factor gives a gaussian onset to equilibrium and multiplies the terms $(1/\rho)(\mathrm{d}\rho/\mathrm{d}z)m_{air}gz$ and $\mathrm{d}P/\mathrm{d}\rho/(r_e+z)$ in Equation (15). A numerical integration of (15) with this factor is shown in **Figure 1**, along with the other solutions given above. All are compared to the standard (measured) atmospheric pressure and density from [17].

Comparing the cyan and blue curves for both pressure and density in **Figure 1**, it is seen that the use of the full generalized equation of hydrostatic equilibrium with Equation (22) arguably gives the best fit to the measured standard atmosphere particularly for the density, and simplifies to the standard adiabatic equation below 15 km, where strict equilibrium does not apply. It further avoids the nonsensical result given by the standard hydrostatic equation, Equation (21), above 30 km.

4. Solution near the Origin

As discussed at the end of Section 1, the unconstrained equilibrium solution gives infinite density at the origin. This defies the Fermi-Dirac equation, which implies a maximum density due to fermion degeneracy pressure. It also seems counter-intuitive, at least in the absence of a gravitational singularity.

To remedy this, a density constraint (as well as a total particle constraint) is added to the Lagrangian density, with a result given by Equation (14). It is not difficult to see that Equation (14) can be written

$$\begin{aligned}&\left\{\mathrm{d}\ln(\rho)/\mathrm{d}\ln(r)+1+2\left[r\lambda_M/\left(\rho m_v^2 G\right)\right]\right\}\delta\rho(r)\\&+\left[\left(4\pi m_v^2 G\right)r\,\mathrm{d}/\mathrm{d}r(\rho/r)\right]\int_r^\infty \mathrm{d}r' r'^2\delta\rho(r')/\left[1+\rho m_v^2 G/(r\lambda_M)\right]\\&\times\left[1+r\lambda_M/\left(\rho m_v^2 G\right)\right]/\left[\mathrm{d}P/\mathrm{d}\rho+m_v M_{enc}(r)G/r+\lambda_\rho\right]=0\end{aligned}\tag{23}$$

The last term of Equation (23), comprising the last two lines, can be estimated in several ways. The simplest approach is to assume that the density variations in the integral average to zero, yielding zero for this term. However, a more careful approach is appropriate, and this can be done assuming the zeroth-order solution, Equation (13). It is further assumed $r\lambda_M/\left(\rho m_v^2 G\right)$ is much less than 1, to investigate behavior near the origin. Then the various factors in the last term of Equation (23) are

$$r\,\mathrm{d}/\mathrm{d}r(\rho/r)=-2\rho_0 r_0/r^2\,,\tag{24a}$$

$$1+r\lambda_M/\left(\rho m_v^2 G\right)=1+O\left(r\lambda_M/\left(\rho m_v^2 G\right)\right),\tag{24b}$$

$$\left[1+\rho m_v^2 G/(r\lambda_M)\right]^{-1}=r^2\lambda_M/\left(\rho_0 r_0 m_v^2 G\right)+o\left(r\lambda_M/\left(\rho m_v^2 G\right)\right),\tag{24c}$$

$$\text{And}\quad m_v M_{enc}(r)G/r=2\pi m_v^2 G\rho_0 r_0 r\equiv C_0 r\,.\tag{24d}$$

The "big-oh" and "little-oh" notations are used in Equations (24b) and (24c). With Equations (24), the last term of Equation (23) can then be estimated as

$$-8\pi\lambda_M\int_r^\infty \mathrm{d}r' r'^2\delta\rho(r')/\left(\mathrm{d}P/\mathrm{d}\rho+C_0 r+\lambda_\rho\right).\tag{25}$$

Note also that with constrained total particle number, the variations should satisfy

$$4\pi\int_0^\infty \mathrm{d}r' r'^2\delta\rho(r')=0\,,\tag{26}$$

which allows Equation (25) to be re-written as

$$+8\pi\lambda_M\int_0^r \mathrm{d}r' r'^2\delta\rho(r')/\left(\mathrm{d}P/\mathrm{d}\rho+C_0 r+\lambda_\rho\right).\tag{27}$$

With bounded variations $\delta\rho(r')$ in 0 to r, this expression tends to zero as r tends to zero, when the denominator is not zero (as should be the case). Hence, however the last term is estimated, the summary result for Equation (23) near the origin is

$$\left\{\mathrm{d}\ln(\rho)/\mathrm{d}\ln(r)+1+2\left[r\lambda_M/\rho m_v^2 G\right]\right\}\delta\rho(r)=0\,.\tag{28}$$

Note that this has two possible solutions for any given r. The first solution is obtained by the usual approach, setting the quantity inside curly brackets to zero. This gives the r^{-1} solution presented in Equation (13) near the origin,

when $r\lambda_M/\left(\rho m_v^2 G\right)$ is much less than 1. A second solution to Equation (28) is obtained by setting the variation $\delta\rho(r)$ to zero, *i.e.*,

$$\rho \text{ = constant in a spherical region.} \tag{29}$$

This latter solution is consistent with an incompressible medium, as occurs when the density is limited by Fermi-Dirac statistics. It is also consistent with the observed features of most celestial bodies, in which the density is approximately constant at their core, and also often in spherical shells. It is also approximately consistent with observations for dark matter in central regions that are less than a kpc in diameter [7] [9]. It implicitly assumes that the properties of the constituent materials are constant within spherical regions. It also is self-consistent with expression (27) equal to zero in such a region. With these facts in mind, the solution (29) of constant density in a region nearest the origin is a candidate solution. It should be further noted that this same result of constant density can be applied far from the origin where particles achieve their nominal cosmic background temperature and density.

Next consider the standard hydrostatic equation. One can assume a solution of the form $\rho_0 r_0 r^{-1}$ for the density and substitute into Equation (1) with Equation (2) and check for consistency. One finds that

$$(1/\rho)(\mathrm{d}P/\mathrm{d}r) = -m_v M_{enc} G/r^2 \to c_\gamma \gamma \rho^{\gamma-2}\,\mathrm{d}\rho/\mathrm{d}r = -2\pi m_v^2 G\rho_0 r_0 . \tag{30}$$

One notes that the left-hand side varies as $r^{-\gamma}$ whereas the right-hand side is a constant under the assumption of $\rho \sim r^{-1}$. Hence for any γ except possibly zero the standard hydrostatic equation is not consistent with an r^{-1} solution.

One can also analyze the standard equation for hydrostatic equilibrium in a gravitational field assuming constant number density ρ_0 in the vicinity of the origin of a spherical mass distribution. A constant density near the origin is self-consistent in Equation (1) as can be seen as follows. Assuming a polytropic relation $P = c_\gamma \rho^\gamma$ in Equation (1) one obtains

$$c_\gamma \gamma \rho^{\gamma-2}\,\mathrm{d}\rho/\mathrm{d}r = -(4\pi/3) m_v^2 G\rho_0 r . \tag{31}$$

here ρ_0 is the sum of OM and DM mass density at the origin, normalized to the mass of the DM particles, with the assumption that OM dominates. This expression shows that the derivative of the dark or ordinary matter is zero at the origin, and gradually becomes more negative, consistent with constant density. A solution of Equation (31) for dark matter gives the following result for $\gamma > 1$:

$$\rho(r) = \left[\rho_{01}^{\gamma-1} - \rho_{01}^{\gamma-1} C_1 r^2\right]^{1/(\gamma-1)} = \rho_{01}\left[1 - C_1 r^2\right]^{1/(\gamma-1)} = \rho_{01}\left[1 - (r/r_a)^2\right]^{1/(\gamma-1)}, \tag{32}$$

where ρ_{01} is the (equivalent) number density at the origin, and $C_1 = (2\pi/3) m_v^2 G\rho_0 (\gamma-1)/\left(c_\gamma \gamma \rho_{01}^{\gamma-1}\right)$. Note that C_1 has units of $\cdot\mathrm{m}^{-2}$, and so one may set $C_1 \equiv r_a^{-2}$, and this is used in the last equality of Equation (32). The solution is clearly approximate because for r greater than r_a the density

can be become complex and also because the density variation of DM in Equation (32) is not included in the enclosed mass in Equation (31). However, it does show that a constant density to order of $(r/r_a)^2$ is consistent with the standard hydrostatic equation.

To summarize Section 4, it is found that setting the variations of the Lagrangian density to zero yields an alternative, constant-density solution. This solution is consistent with the observed density profile of many celestial bodies near their origin, including observationally-inferred dark matter profiles. The standard hydrostatic equation also shows solutions consistent with approximately constant density at the origin but is seen to be inconsistent with r^{-1} solutions. With these considerations in mind, the constant-density solution is used in a region of radius r_0 near the origin in the following sections.

5. Behavior of Solutions away from the Origin

The mass-constrained Equation (14) is most relevant here. The last term in the equation is assumed negligible in this case, since the integral of any physical density variations from r to infinity should tend to zero for large r in the case of constrained total particle number. The result is

$$\begin{aligned}&\left[1+r\lambda_M/\left(\rho m_v^2 G\right)\right]^{-1}\left[(1/\rho)\mathrm{d}\rho/\mathrm{d}r\right]\left[\mathrm{d}P/\mathrm{d}\rho+m_v M_{enc}(r)G/r+\lambda_\rho\right]\\&+\left\{2-\left[1+r\lambda_M/\rho m_v^2 G\right]^{-1}\right\}\left[\mathrm{d}P/\mathrm{d}\rho+m_v M_{enc}(r)G/r+\lambda_\rho\right]/r=0.\end{aligned}\tag{33}$$

Inspection shows that this equation can be simplified considerably, giving, as in Equation (28),

$$\mathrm{d}\ln(\rho)/\mathrm{d}\ln(r)=-1-2\left[r\lambda_M/\rho m^2 G\right].\tag{34}$$

This result shows that the solution is independent of the density constraint λ_ρ with the neglect of the last term of Equation (14). This equation may further be put into dimensionless form by defining $\lambda_M/\left(m_v^2 G\right)\equiv\rho_c/r_c$, where is ρ_c the cutoff density and r_c is the cutoff radius where the density drops due to the constraint on the total particle count. The resulting equation is

$$\mathrm{d}\ln(\rho)/\mathrm{d}\ln(r)=-1-2\left[(r/r_c)(\rho_c/\rho)\right].\tag{35}$$

This equation may be solved numerically or perturbatively. For a perturbative treatment, assume $(r/r_c)(\rho_c/\rho)$ is much less than 1, corresponding to radii near the origin but not at the origin. Then the zeroth order solution is given by Equation (12),

$$\rho(r)=\rho_0(r_0/r)+O\left[(r/r_c)(\rho_c/\rho)\right].\tag{36}$$

where the "big-oh" notation is used to indicate the order of the approximation. Substituting this solution back into Equation (35) gives

$$\mathrm{d}\ln(\rho)/\mathrm{d}\ln(r)=-1-2\left[\left(r^2/r_c r_0\right)\left(\rho_c/\rho_0\right)\right].\tag{37}$$

This equation is readily solved to give

$$\begin{aligned}\rho(r) &= \rho_0(r_0/r)\exp\left\{-\left(r^2-r_0^2\right)\left[\rho_c/(\rho_0 r_0 r_c)\right]\right\} + O\left\{\left[(r/r_c)(\rho_c/\rho)\right]^2\right\} \\ &= \rho_0(r_0/r)\left\{1-\left(r^2-r_0^2\right)\left[\rho_c/(\rho_0 r_0 r_c)\right]\right\} + O\left\{\left[(r/r_c)(\rho_c/\rho)\right]^2\right\} \\ &= \frac{\rho_0 r_0}{r\left[1+\left(\frac{r^2}{r_0^2}-1\right)\left(\frac{\rho_c r_0}{\rho_0 r_c}\right)\right]} + O\left\{\left[(r/r_c)(\rho_c/\rho)\right]^2\right\} \\ &= \frac{\rho_0 r_{01}}{r\left(1+r^2/r_{02}^2\right)} + O\left\{\left[(r/r_c)(\rho_c/\rho)\right]^2\right\}\end{aligned} \tag{38}$$

In the last expression, $r_{01} = r_0/(1-\mu)$, $r_{02}^2 = r_0^2(1-\mu)/\mu$, and $\mu = \rho_c r_0/(\rho_0 r_c)$. With μ less than 1 and $\left[(r/r_c)(\rho_c/\rho)\right]^2 < 1$, this expression is similar to the NFW result with a $1/r$ dependence near the origin and a $1/r^3$ dependence far from the origin. It is also similar to the double-power-law models and in particular the (α,β,γ) models [2] [3], with $(\alpha,\beta,\gamma) = (2,3,1)$. However, because the scaling factor r_{02} differs from r_{01}, the above is formally different from such models.

To check the above results, Equation (35) may be integrated numerically from the origin for some relevant input values. In accord with Section 4, the density is assumed constant for $r < r_0$. The resulting density is shown in plot (a) of **Figure 2**, along with approximate results from Equations (36) and (38). The dimensionless inputs for **Figure 2** are $r_0/r_c = 0.011$ and $\rho_c/\rho_0 = 0.011$. One can see that the approximation of Equation (36) is relatively good for r less than about $0.5r_c$ and for Equation (38) for r less than about $0.7r_c$. The plot on the right shows the normalized enclosed-mass curve for the numerical result in plot (a). The normalization of enclosed mass uses the results of Section 4, accounting for the region of radius r_0 of constant density ρ_0 near the origin and a $1/r$ dependence outside this region to a radius of $0.7r_c$ (at which the $1/r$ dependence is seen to be no longer valid). The result is

$$M_{enc,tot} \equiv 2\pi m_v \rho_0 r_0 \left[(0.7r_c)^2 - r_0^2/3\right]. \tag{39}$$

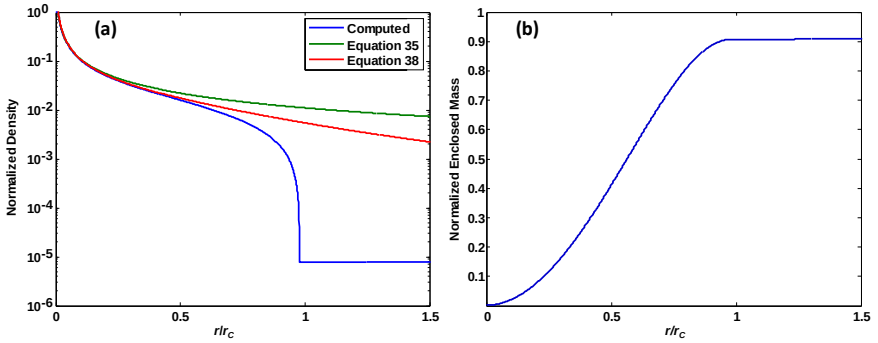


Figure 2. (a): Normalized DM density ρ/ρ_0 versus normalized radius. Blue: numerically computed from Equation (35), thresholded from below at $10^{-5}\rho_0$. Green: leading order result, Equation (36). Red: next leading order result, Equation (38). (b): normalized enclosed mass $M_{enc}(r)/M_{enc,tot}$ versus normalized radius for computed density profile.

It should be noted that the results of Figure 2 are not explicitly dependent on the polytropic exponent, particle mass, or temperature. The result does depend on the latter 2 parameters implicitly through the initial density at or near the origin via Equation (4), for example. One may observe from Equation (35) that the logarithmic slope varies continuously from −1 at the origin to −3 for $r/\rho = r_c/\rho_c$. It is well known that logarithmic exponents less than −3 correspond to solutions that have convergent enclosed mass [1] [2].

In both this section and Section 4, the last term of Equation (14) is neglected with some justification. However, there is no particular justification for its neglect in regions neither near nor far from the origin. The retention of this last term could introduce perturbations into the nominal equation which are arbitrary, but which might average in some sense to zero. It should also be observed that in cases of interest, the parameter λ_M is extremely small compared to the other energy densities involved, so in such cases the neglect of the last term of Equation (14) is justified. For example, with $\lambda_M = m_v^2 G\rho_c/r_c$, with r_c = 92 kpc for a halo like that of the Milky Way [19], a temperature of 2 K and masses of 1 eV/c^2 and 5 keV/c^2, one obtains $\lambda_M = 2.13\times10^{-90}$ and 1.89×10^{-77} J·m^{-3}, respectively, using Equation (4). These values of λ_M should be compared with the pressure at the origin, which from Equations (4) and (5) is about 3.22×10^{-10} and 1.14×10^{-4} J·m^{-3} for the same two masses, respectively. Even with densities 200 times lower as occur might occur at the edge of the density profile, the value of λ_M is negligible in comparison. These considerations provide some justification for neglect of the last term of Equation (14).

6. Comparison to Einasto and De-Projected Sersic Profiles

Taking all the above sections into consideration, one finds that $r\lambda_M/\left(\rho m_v^2 G\right) = \left(r/r_c\right)\left(\rho_c/\rho\right)$ is always less than 1 where there is appreciable density, so the resulting solution is

$$\rho(r) = \begin{cases} \rho_0 \text{ in a spherical region about the origin of radius } r_0, \text{or} \\ \rho(r) \text{ satisfies } \dfrac{\mathrm{d}\ln\rho}{\mathrm{d}\ln r} = -1 - 2\left(\dfrac{r}{r_c}\right)\left(\dfrac{\rho_c}{\rho}\right) \end{cases} \tag{40}$$

This solution has a central core region that is qualitatively consistent with the size of inner profiles of low-spectral brightness galaxies [2] [8] [9]. Further, the constant density in a core region avoids a density singularity and so is consistent with a Fermi-Dirac density distribution at finite temperature. Moreover, the finite density at the origin is consistent with all observed astronomical bodies except black holes. Lastly, the "core" solution offers a potential path to resolution to the "core-cusp" problem in dwarf galaxies.

For the region just outside the core, the logarithmic slope given by Equation (40) is −1 or less, which is consistent with most past results. A comparison of the solution to the $(\alpha,\beta,\gamma) = (2,3,1)$ model is given in the previous section. Equation (40) can also be compared to the well-known de-projected Sersic and Ei-

nasto models [4] [5] [6]. The de-projected Sersic (dpS) density distribution for radiant matter is approximated by [2]:

$$\rho_{rad} = \rho_{rad0}\left(r/R_e\right)^{-p_n}\exp\left[-b_n\left(r/R_e\right)^{1/n}\right]. \tag{41}$$

The parameter ρ_{rad0} is obtained by setting the volume integral of Equation (41) equal to the measured or inferred ordinary mass of the galaxy. The variable R_e denotes the radius which encloses 1/2 the total light of the galaxy. The dpS profile is also used with some success for characterization of DM density profiles versus radius. The other two parameters in Equation (41) are given conveniently and approximately from Equations (19) and (27) of [2],

$$p_n = 1.0 - 0.6097/n + 0.05463/n^2 \text{ , and} \tag{42}$$

$$b_n = 2n - 1/3 + 0.009876/n \text{ .} \tag{43}$$

The parameter n ranges from roughly 2 to 4 for DM haloes of galaxies and clusters, from the same reference. With these values of n, the leading exponent p_n in Equation (42) is −0.7 to −0.85, which is not too different from −1, and [2] notes that the logarithmic slope is almost always approximated by −1 in the vicinity of the origin in the N-body simulations they investigated. The logarithmic derivative of Equation (41) is

$$\mathrm{d}\ln\left(\rho_{dpS}\right)/\mathrm{d}\ln\left(r\right) = -p_n - \left(b_n/n\right)\left(r/R_e\right)^{1/n}. \tag{44}$$

Another density distribution used for characterization of DM is the Einasto distribution, which for the purposes of this paper is given by

$$\rho_{Ein} = \rho_0\exp\left[-d_n\left(r/r_e\right)^{1/n}\right], \tag{45}$$

where r_e is the $\exp\left(-d_n\right)$ radius of the DM galactic halo, n is typically 4 to 7 in this case, and d_n is given approximately by Equation (24) of [2]:

$$d_n \approx 3n - 1/3 + 0.0079/n, \text{ for } n > 0.5 \text{ .} \tag{46}$$

For the Einasto distribution, the logarithmic derivative is

$$\mathrm{d}\ln\left(\rho_{Ein}\right)/\mathrm{d}\ln\left(r\right) = -\left(d_n/n\right)\left(r/r_e\right)^{1/n}. \tag{47}$$

Inspection of Equation (44) finds a total logarithmic slope of about $-3 + 1/\left(3n\right) + 0.6097/n$ at $r = R_e$ for the dpS model. The leading term matches the logarithmic slope of Equation (40) at $\left(r/r_c\right)\left(\rho_c/\rho\right) = 1$. Note that the onset of steeper logarithmic slopes occurs for smaller r with the dpS model than for Equation (40). The primary potential disadvantage of the dpS model relative to Equation (40) is the infinite density at the origin (but if there is a gravitational singularity this may be an advantage).

The Einasto model does not have infinite density at the origin. The logarithmic slope of the Einasto model is also zero at the origin, matching the core solution of Equation (40). These may be viewed as positive features if the true density distribution is not infinite at the origin. However, the derivative $\mathrm{d}\rho_{Ein}/\mathrm{d}r$ is infinite at the origin for $n > 1$. The logarithmic slope of the Einasto model also

gives a value of about −3 + 1/(3*n*) at $r = r_e$ using Equation (46). This is an approximate match of the logarithmic slope of Equation (40) at $(r/r_c)(\rho_c/\rho) = 1$. However, as with the dpS model, the onset of steeper logarithmic slopes, slopes less than −1 but greater than −3, occurs for smaller *r* than for Equation (40).

This different behavior of the model solutions can be addressed with the possible solutions or extensions of Equation (14). One approach is to add angular momentum. The net effect of angular momentum would be to limit the extent of the galactic halo and would result in a more rapid decline in density than given in Equation (40). However, faster declines are found in simulations without appeal to angular momentum. As a second approach, one may consider the results of Section 3, which indicate that if the medium is not in a strict equilibrium, then solutions with faster declines occur. The outer perimeters of a large halo are not expected to be in strict equilibrium because of the process of aggregation of subhaloes [10]. Proceeding as in Section 3, one may neglect the term $(1/\rho)(\mathrm{d}\rho/\mathrm{d}r)\left[m_v M_{enc}(r) G/r\right]$ in the first line of Equation (14) and the $(\mathrm{d}P/\mathrm{d}\rho)/r$ term in the second line, as well as the last term with leading factor λ_M. The resulting equation is

$$(1/\rho)\mathrm{d}P/\mathrm{d}r = -m_v M_{enc}(r) G/r^2 \left[1 + 2(r/r_c)(\rho_c/\rho)\right]. \tag{48}$$

Consider the region *r* which is significantly greater than r_0 but also significantly less than $r_c(\rho/\rho_c)$, to address the issue identified in the previous paragraph. Using the techniques of Equations (30) to (32), one may assume an initial $\rho_0(r_0/r)$ solution to evaluate the enclosed mass, but this time without the expectation for consistency of the resulting density profile. One obtains an equation similar to Equation (21) in this range of *r*:

$$\begin{aligned} \rho(r) &= \rho_0 \left[1 - r/r_\gamma\right]^{1/(\gamma-1)} \\ &= \rho_0 \left[1 + r/r_\gamma\right]^{-1/\gamma-1} + o(r/r_\gamma) + o\left[(r/r_c)(\rho_c/\rho)\right] \end{aligned} \tag{49}$$

where

$$r_\gamma = c_\gamma \gamma / \left[2\pi(\gamma - 1) G m_v^2 r_0 \rho_0^{2-\gamma}\right]. \tag{50}$$

For related states of matter, such as neutron stars, a polytropic exponent γ ranging from 3/2 to 2 is often used [20] [21] [22]. With these choices for γ, one obtains familiar forms of solutions, $\left[1 + r/r_{3/2}\right]^{-2}$ to $\left[1 + r/r_2\right]^{-1}$, respectively. The former can be found as a factor in the NFW solution. The form of these solutions, along with the large values of r_γ found for certain particle masses and temperatures as shown in Section 7, provides a potential explanation for the long tails observed in the Einasto and dpS model profiles. Also, in this range of *r* one may see that $s = 2/(2 - \gamma)$ can give self-consistent solutions in Equation (48) by inserting $\rho \sim 1/r^s$, for $s > 0$.

One may also compare the logarithmic derivative of Equation (48) to the dpS and Einasto forms. From Equation (48), the logarithmic derivative is given by

$$\mathrm{d}\ln(\rho)/\mathrm{d}\ln(r) = -m_v M_{enc}(r) G/r \left[1 + 2(r/r_c)(\rho_c/\rho)\right] / c_\gamma \gamma \rho^{\gamma-1}. \tag{51}$$

Substituting $\rho \sim 1/r^s$, $s > 0$, gives a right-hand side proportional to $r^{2+s(\gamma-2)}$ for $(r/r_c)(\rho_c/\rho) < 0.5$. This is always negative (with the overall minus sign on the right of Equation (51)). When the right-hand side is less than −1, the density will decline faster than the r^{-1} solution. One can match the $r^{1/n}$ dependence of the dpS and Einasto logarithmic derivatives of Equations (44) and (47) with exponent n of 2 to 7, by setting $2 + s(\gamma - 2)$ equal to n^{-1}. The requisite values of s are in the range of $(13/7)/(2-\gamma)$ to $(3/2)/(2-\gamma)$, for $\gamma < 2$ and in the stated range of values of r. Hence, the overall radial dependence of the Einasto and de-projected Sersic models can be obtained with the assumption of deviation from equilibrium, using insights from Section 3. One also observes that including the mass constraint term will invariably result in convergent mass for decreasing density with increasing radius.

7. Lane-Emden Solutions for Non-Relativistic Particles

It is also instructive to consider the case in which constituent particles are non-relativistic and use the standard hydrostatic relation Equation (1) and the polytropic relation Equation (2). This leads to a familiar equation for the density and provides an analytic estimate for the scale size of cosmic structures. When Equation (2) for non-relativistic fermions is inserted in the standard equation for hydrostatic equilibrium, Equation (1), one obtains

$$(5/3)c_{5/3}\rho^{-1/3}\,\mathrm{d}\rho/\mathrm{d}r = -m_v M_{enc}(r)G/r^2 . \tag{52}$$

Multiplying Equation (52) by r^2, and then including the $\rho^{-1/3}$ factor in the differential, one obtains

$$(5/2)r^2 c_{5/3}\,\mathrm{d}\rho^{2/3}/\mathrm{d}r = -m_v M_{enc}(r)G . \tag{53}$$

Dividing by $m_\upsilon^2 G$ and applying the operator $r^{-2}\,\mathrm{d}/\mathrm{d}r$ to both sides of Equation (53) gives

$$\begin{aligned}&\left[5c_{5/3}/\left(2m_\upsilon^2 G\right)\right]r^{-2}\,\mathrm{d}/\mathrm{d}r\left(r^2\,\mathrm{d}\rho^{5/3}/\mathrm{d}r\right)\\ &= C_{5/3}\nabla^2\rho^{2/3} = -r^{-2}\,\mathrm{d}/\mathrm{d}r\,M_{enc}(r)/m_\upsilon\end{aligned} \tag{54}$$

where $C_{5/3} = 5c_{5/3}/\left(2m_\upsilon^2 G\right)$. Note that the first equality in Equation (54) assumes spherical symmetry. Account for ordinary radiant matter as part of the enclosed mass. Then

$$M_{enc}(r) = 4\pi\int_0^r \mathrm{d}r' r'^2\left[m_v\rho(r') + \rho_{m,rad}(r')\right], \tag{55}$$

where $\rho_{m,rad}(r)$ is the density of ordinary matter, as defined in Section 2. Also normalize the number density by its value ρ_0 at the origin. With the notatio $\rho_1 = \rho/\rho_0$ the resulting dimensionless equation is

$$\left[5c_{5/3}/\left(8\pi m_\upsilon^2 G\rho_0^{1/3}\right)\right]\nabla^2\rho_1^{2/3} = -\rho_1 - \rho_{m,rad}/\left(m_v\rho_0\right). \tag{56}$$

This equation is a form of the well-known Lane-Emden equation for polytropic spheres [13] for a polytropic exponent of 5/3. The constant

$$L_{5/3} = \left[5c_{5/3} \Big/ \left(8\pi m_\upsilon^2 G \rho_0^{1/3} \right) \right]^{1/2} \quad \text{(non-relativistic)}. \tag{57}$$

has units of length and sets the scale of the decay of the density distribution. For relativistic fermions with polytropic exponent 4/3 the corresponding length scale is

$$L_{4/3} = \left[c_{4/3} \Big/ \left(\pi m_\upsilon^2 G \rho_0^{2/3} \right) \right]^{1/2} \quad \text{(relativistic)}. \tag{58}$$

The parameter $L_{5/3}$ is shown in **Table 1** as a function of particle mass for temperatures of T = 100, 8, and 2 K respectively at the origin. These temperatures set the density ρ_0 at the origin via Equation (4) for fermions. The temperatures larger than 2 K may be relevant for bound fermion states, for which the chemical potential, Fermi energy, and corresponding Fermi temperature may be much greater than ambient. Also shown is $r_{5/3}$ from Equation (50). The scale sizes shown for larger temperatures are much smaller because a larger temperature implies a higher density, which therefore results in a smaller scale parameter due to stronger gravitational forces from more mass. To understand the mass scaling, recall that number density scales as (particle mass)$^{3/2}$ for non-relativistic fermions [23], so larger masses also result in larger density which further results in smaller scale sizes. Accounting for this number density scaling and the scaling of $c_{5/3}$ with mass, $L_{5/3}$ scales as $m_\upsilon^{-7/4}$ and $r_{5/3}$ scales as $m_\upsilon^{-7/2}$. **Table 1** shows that the lower masses, ~0.025 eV/c^2, and lower temperature, 2 to 8 K, yield values of the scale parameter $L_{5/3}$ which correspond to the peak of the large-scale mass power spectrum of about 300 Mpc [24]. Somewhat higher masses, ~0.1 eV/c^2 and temperatures of 2 to 8 K correspond to scales sizes $L_{5/3}$ of superclusters of the order of 30 Mpc [25] [26]. Masses of 20 eV/c^2 or more for these temperatures give $L_{5/3}$ of 3 kpc or less, corresponding to the diameter of dwarf galaxies. One may note that the temperatures and masses shown correspond to non-relativistic conditions.

Table 1. Scale parameters $L_{5/3}$ and $r_{5/3}$ (Mpc) versus T and particle mass.

$L_{5/3}$	Mass (eV/c^2)						
T (K)	0.025	0.1	0.4	1.6	5.0	10	20
2	391	34.7	3.07	0.27	0.037	0.011	0.0033
8	275	24.5	2.17	0.192	0.026	0.0078	0.0023
100	133	12.7	1.14	0.102	0.0139	0.0041	0.0006
$r_{5/3}$	Mass (eV/c^2)						
T (K)	0.025	0.1	0.4	1.6	5.0	10	20
2	3.1×10^8	2.4×10^6	1.9×10^4	147	2.72	0.240	0.021
8	1.5×10^8	1.2×10^6	9390	73.4	1.36	0.120	0.011
100	3.6×10^7	3.2×10^5	2620	20.7	0.385	0.033	0.003

Table 1 shows results for $r_{5/3}$, assuming r_0 = 1 kpc. The table highlights the rapid variability of this scale parameter with mass. Masses of 1.2 to 1.6 eV/c^2 give scale sizes corresponding to the peak of the large-scale mass power spectrum, at a nominal background temperature of 2 K, and masses of 10 to 50 eV/c^2 yield scale sizes that correspond to radii of large to dwarf galaxies. Note that the masses shown in this table are not consistent with recent estimates of dark-matter particle mass [27] [28], which identify masses of the order of 5 keV. This implies that some underlying assumption of this paper is not consistent with the assumptions of those results. The only key assumptions of Table 1 are (a) mechanical or energy equilibrium applies, (b) the particles are fermions, and (c) ordinary gravity applies. The value of $r_{5/3}$ also depends on r_0, the region in which the density is approximately constant and equal to ρ_0. This region is a corollary of the assumption of fermionic matter in the absence of a gravitational singularity.

Figure 3 shows some density profiles corresponding to Table 1. Plot (a) of Figure 3 shows results for the Lane-Emden equation, Equation (56). Plot (b) of Figure 3 shows results for the generalized hydrostatic equation, Equation (40). The assumed conditions are shown in Table 2. For the numerical integration, 4000 radial steps are used, with each step equal to 77 kpc. The seed mass at $r = 0$ is set to the estimated mass of OM of a galaxy similar to the Milky Way, which is chosen to be 9×10^{10} solar masses [19]. The chemical potential is set to zero for the results of Figure 3 because the particles are assumed to be free and non-interacting rather than bound (however, the results shown in plot (b) of Figure 3 are also consistent with bound states of matter in fluid or gaseous form). The temperature T at the origin for plot (b) of Figure 3 is set to 25 K in order to remain above a number density floor of 10^8 m^{-3} out to a radius r_c of 300 Mpc. This choice of number density floor is based on the estimated neutrino density in standard cosmology [29]. The average temperature of the profiles on the right is about 0.15 K in all cases, using Equation (4) to relate density to temperature. Note also that plot (b) of Figure 3 derives the r^{-1} solution. For this solution, the scale parameter r_γ of Equation (50) and Table 1 is not directly relevant.

Table 2. Inputs for Figure 3.

Input Parameter	Values	Comment
Temperature at origin, T	2, 25 K	Sets density at origin using Equation (4), $\mu_\upsilon = 0$.
Masses of particles, m_υ	0.025 - 0.3 eV/c^2	
Source mass at origin	9×10^{10} solar masses	Similar to Milky Way OM mass
Inner scale r_0	77 kpc	Input to Equation (40)
Temperature T_{out} at outer radius r_c	0.1 K	Sets density ρ_c at outer radius in Equation (40) using Equation (4).
Polytropic exponent	5/3	Input to Lane-Emden equation

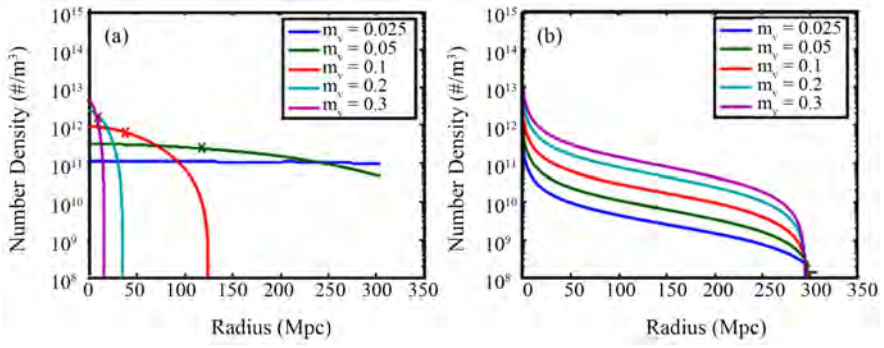


Figure 3. Particle density versus radius (Mpc). (a) Solution to Lane-Emden equation, Equation (56), with polytropic exponent = 5/3 and $T = 2$ K. (b): Solution to generalized hydrostatic Equation (40) with $T = 25$ K, $r_c = 300$ Mpc. In (a), the X's mark the corresponding large-scale structure radius $L_{5/3}$ from Equation (57).

In summary, this Section compares the results of the Lane-Emden equation with that of the generalized hydrostatic equation. It is seen that the former produces a different dependence of scale size on temperature and particle mass that does the generalized hydrostatic equation, and the shapes of the solutions are significantly different near the origin; the former is flat, the latter has a cusp at or near the origin. However, in both cases, a narrow range of masses and temperatures yield scale sizes that are consistent with observed galactic and large-scale structure as seen from Table 1.

8. Summary

In summary, Section 2 shows a derivation of a generalized hydrostatic equation in spherically symmetric geometries with a single species of matter satisfying a polytropic relation. Section 3 provides a partial validation of the generalized hydrostatic Equation (10) using the earth's atmosphere. Section 4 shows that there is a constant-density solution for the Lagrangian formulation in a spherical region about the origin (or in spherical regions about the origin), and that the r^{-1} solution near the origin is not self-consistent with the standard hydrostatic equation, as expected, when the added terms of the generalized equation are not included. Section 5 explains how Equation (14) gives a convergent enclosed mass far from the origin and shows an example solution of the differential equation. Section 6 demonstrates that the solutions can match the radial behavior of accepted models of OM and DM density versus radius away from the origin. Section 7 presents solutions of the generalized equation for the density and compares it to the standard hydrostatic equation with a polytropic exponent in a gravitational potential. The overall conclusion is that a generalized equation for hydrostatic equilibrium, Equation (40), gives a cuspy solution for DM that is more consistent with observations and simulations (e.g., [1] [2]) than the standard hydrostatic equation near the origin and more consistent with finite density and lower slope in the immediate vicinity of the origin [7] [9] than the de-projected Sersic or Einasto models. Because this paper shows that the r^{-1} den-

sity profile is largely independent of the properties of the constituent material, this profile should indeed be "ubiquitous," as discussed by other authors, wherever there is strict equilibrium.

Portions of this work were presented in Paper APR19-000356 at the 2019 April Meeting of the American Physical Society.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Navarro, J.F., Frenk, C.S. and White, S.D.M. (1996) *The Astrophysical Journal*, **462**, 563-575. https://doi.org/10.1086/177173

[2] Merritt, D., Graham, A.W., Moore, B., Diemand, J. and Terzic, B. (2006) *The Astronomical Journal*, **132**, 2685-2700. https://doi.org/10.1086/508988

[3] Hernquist, L. (1990) *The Astrophysical Journal*, **356**, 359-364. https://doi.org/10.1086/168845

[4] Sérsic, J.L. (1963) *Boletín de la Asociación Argentina de Astronomía*, **6**, 41-43.

[5] Prugniel, P. and Simien, F. (1997) *Astronomy and Astrophysics*, **321**, 111-122.

[6] Einasto, J. (1965) *Alma-Ata*, **5**, 87.

[7] Graham, A., Merritt, D., Moore, B., Diemand, J. and Terzic, B. (2006) *The Astronomical Journal*, **132**, 2701-2710. https://doi.org/10.1086/508990

[8] de Blok, W.J.G. (2010) *Advances in Astronomy*, **2010**, Article ID: 789293. https://doi.org/10.1155/2010/789293

[9] Oh, S.-H., Hunter, D.A., Brinks, E., *et al.* (2015) *The Astronomical Journal*, **149**, 180-276. https://doi.org/10.1088/0004-6256/149/6/180

[10] Zavala, J. and Frenk, C.S. (2019) *Galaxies*, **7**, 81-135. https://doi.org/10.3390/galaxies7040081

[11] Robertson, A., Massey, R. and Eke, V. (2017) *Monthly Notices of the Royal Astronomical Society*, **465**, 569-587. https://doi.org/10.1093/mnras/stw2670

[12] Ringwald, A. and Wong, Y.Y.Y. (2004) *Journal of Cosmology and Astrophysics*, **12**, 005. https://doi.org/10.1088/1475-7516/2004/12/005

[13] Lane, J.H. (1870) *American Journal of Science*, **2**, 57-74. https://doi.org/10.2475/ajs.s2-50.148.57

[14] Graham, A.W. and Driver, S.P. (2005) *Publications of the Astronomical Society of Australia*, **22**, 118-127. https://doi.org/10.1071/AS05001

[15] Graham, A.W. and Spitler, L. (2009) *Monthly Notices of the Royal Astronomical Society*, **397**, 2148-2162. https://doi.org/10.1111/j.1365-2966.2009.15118.x

[16] Weinzirl, T., Jogee, S., Khochfar, S., Burkert, A. and Kormendy, J. (2009) *The Astrophysical Journal*, **696**, 411-447. https://doi.org/10.1088/0004-637X/696/1/411

[17] White, F.M. (1986) Fluid Mechanics. McGraw-Hill, New York.

[18] ICAO (1993) Manual of the ICAO Standard Atmosphere. 3rd Edition, Montreal.

[19] Kafle, P.R., Sharma, S., Lewis, G.F. and Bland-Hawthorn, J. (2014) *The Astrophysical Journal*, **794**, 59. https://doi.org/10.1088/0004-637X/794/1/59

[20] Ruffert, M. and Janka, H.-T. (2010) *Astronomy and Astrophysics*, **514**, A66.

https://doi.org/10.1051/0004-6361/200912738

[21] Cho, H.S. and Lee, C.H. (2010) *Publications of the Astronomical Society of Japan*, **62**, 315-321. https://doi.org/10.1093/pasj/62.2.315

[22] Ozel, F. and Freire, P. (2016) *Annual Reviews of Astronomy and Astrophysics*, **54**, 401-440. https://doi.org/10.1146/annurev-astro-081915-023322

[23] Koester, D. and Chanmugam, G. (1990) *Reports on Progress in Physics*, **53**, 837-915. https://doi.org/10.1088/0034-4885/53/7/001

[24] Percival, W.J., Nichol, R.C., Eisenstein, D.J., Frieman, J.A., Fukugita, M., *et al.* (2007) *The Astrophysical Journal*, **657**, 645-663. https://doi.org/10.1086/510615

[25] Gregory, S.A. and Thompson, L.A. (1978) *The Astrophysical Journal*, **222**, 784-799. https://doi.org/10.1086/156198

[26] Oort, J.H. (1983) *Annual Reviews Astronomy and Astrophysics*, **21**, 373-428. https://doi.org/10.1146/annurev.aa.21.090183.002105

[27] Iršič, V., Viel, M., Haehnelt, M.G., Bolton, J.S., Cristiani, S., *et al.* (2017) *Physical Review D*, **96**, Article ID: 023522. https://doi.org/10.1103/PhysRevD.96.023522

[28] Hsueh, J.-W., Enzi, W., Vegetti, S., Auger, M.W., Fassnacht, C.D., *et al.* (2019) *Monthly Notices of the Royal Astronomical Society*, **492**, 3047-3059. https://doi.org/10.1093/mnras/stz3177

[29] Perkins, D.H. (2000) Introduction to High Energy Physics. 4th Edition, Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511809040

Journal of Modern Physics, 2020, 11, 668-682
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Photon Can Be Described as the Normalized Mutual Energy Flow

Shuang-Ren Zhao

mutualenergy.org, London, Canada
Email: shrzhao@gmail.com

How to cite this paper: Zhao, S.-R. (2020) Photon Can Be Described as the Normalized Mutual Energy Flow. *Journal of Modern Physics*, **11**, 668-682.
https://doi.org/10.4236/jmp.2020.115043

Received: March 28, 2020
Accepted: May 5, 2020
Published: May 8, 2020

Abstract

Einstein guessed that the macroscopic electromagnetic wave is built by thousands of photons, however, no one has offered a theory about how the macroscopic electromagnetic wave is built from photons. A concrete theory about photons is needed to answer this question. Current theory for photons is Maxwell's equation which has the solution of waves, but it is difficult to describe the photon as a particle. There is the paradox problem of wave-particle duality. This article offers one solution to solve this problem by introducing the normalized mutual energy flow. The interaction of the retarded wave and advanced wave produce the mutual energy flow. The mutual energy flow satisfies the mutual energy flow theorem. The mutual energy flow theorem tells us that the energy that goes through each surface between the emitter and the absorber is all same. That means the mutual energy flow is different in comparison to the waves. The wave, for example, the retarded wave, its amplitude is decreased with the distance from the source to the point of the field. The mutual energy flow does not decrease. The author noticed this and claimed that the photon is the mutual energy flow. In this article the author updated this claim that the photon is the normalized mutual energy flow. Here the normalization of mutual energy flow will normalize the mutual energy flow to the energy of a photon, which is E = hf. E is the energy of the photon; h is Planck constant; f is the frequency of the light. This normalization is similar to the normalization in quantum mechanics. After this normalization the relation between an electromagnetic wave and photon as a particle becomes clear. This article will prove that the macroscopic wave of an electromagnetic field can be built by thousands of normalized mutual energy flows, which describes the photons. The mutual energy flow is an interaction of the retarded wave and the advanced wave. The retarded wave and the advanced wave satisfy the Maxwell equations. There are two additional waves which are the time-reversal waves which satisfy time-reversal Maxwell equations. The advanced wave and the two time-reversal waves are all real and physical electromagnetic fields. The time-reversal waves cancel all self-energy

flows of the retarded wave and advanced wave. Hence, the waves do not carry any energy, the energy is only transferred by the normalized mutual energy flows which are the photons. Hence, all energy is transferred by the photon instead of waves. This offers a solution to paradox of the duality of wave-particle.

Keywords

Advanced Wave, Retarded Wave, Time-Reversal Wave, Photon, Mutual Energy, Energy Flow, Electromagnetic Fields, Normalization, Wave-Particle Duality, Electron, Quantum

1. Introduction

1.1. Mutual Energy Flow Can Be Applied to Normalize the Photon

In quantum mechanics we can make normalization to an electron in the orbiter, which belongs to the stable solution of Schrödinger equation. We also can normalize the plane waves. However, we do not know how the normalization can be done for photon, or the electron in empty space. The reason is that we do not know how to describe the photon or electron in empty space. Someone perhaps will argue the photon can be described by Maxwell equations in the empty space, the electron can also be described as Schrödinger equation in empty space. That is true, but photon starts from an emitter and ends at an absorber. The energy is transferred from a point to another point without decrease. The solution of Maxwell equation is a wave which spreads to the whole space; hence, the amplitude of the wave decreases with the distance from the field point to the source. The wave solution and the particle have very big difference. Someone perhaps will argue that plane wave can describe the particle. But first, the charge cannot radiate plane wave, second the plane wave doesn't focus to two points in the place the photon is radiated and received. This is so called wave and particle duality paradox. Since this paradox, how the normalization of the photon in empty space occurs has not been solved yet. Similar situation happens also for electron in empty space. We need a wave package which can focus to point in both places of emitter and absorber; it is thick in the middle between the emitter and the absorber.

Recently this author has claimed that the photon is nothing else but the mutual energy flow [1]. The mutual energy flow is the interaction between the retarded wave sent from the source and the advanced wave sent from the absorber. This integration leads to an energy flow that starts from emitter ends at the absorber. The mutual energy flow theorem guarantees the energies go through any surfaces between the emitter and the absorber are all equal. Hence, the mutual energy flow can be a good role to do the normalization for photon. This article describes the details why and how to normalize the mutual energy flow and claim the photon is normalized mutual energy flow.

1.2. History Review about the Theory for Mutual Energy Flow

In physics, Wheeler and Feynman introduced the absorber theory in 1945 and 1949 [2] [3], which claimed that the electric current can radiate advanced wave in the same time when it radiates the retarded wave. The absorbers send the advanced waves. This view of point has been extended to the transactional interpretation of quantum mechanics by Cramer around 1980 [4]. The theory about advanced wave of Wheeler and Feynman cited the earliest publications of action at a distance [5] [6] [7].

In electronic engineering, Welch has introduced his time-domain reciprocity theorem in 1960 [8]. In this theorem the transmitting antenna sends the retarded wave. The receiving antenna sends the advanced wave. Ramsey introduces his reciprocity theorem in 1963 [9]. This author introduced the mutual energy theorem in early of 1987 [10]. de Hoop introduced the cross-correlation reciprocity theorem in the end of 1987 [11]. All these theorems are same theorem and only have a difference in Fourier transform. Traditionally, when we speak about the reciprocity theorem, the two fields in the theorem do not need both to really exist. One field can be real and another can be virtual. In the reciprocity theorem of Welch, the retarded wave can be real, the advanced wave can be virtual, hence, the causality of physics has not been violated. However, when we speak these theorems as an energy theorem, it needs the two fields in the theorem to really exist. Hence, this author claimed this theorem is an energy theorem, that means the advanced wave must also exist in nature as a real substance. This author also applied the mutual energy theorem to Huygens principle, sphere wave and plane wave expansions in 1989 [12] [13].

Recently, this author reviewed the mutual energy theorem. Since, the search engine technology is much developed than before, when this time entering this field, this author found the cross-correlated reciprocity theorem of de Hoop [11]. From that, the author further found the time-domain reciprocity theorem of Welch [8]. Welch talked the concept of advanced wave, this author began to search the concept of advanced wave and noticed the publications of Wheeler and Feynman [2] [3], Cramer [4]. This author also noticed the publication of Stephenson [14], in which has a lot of interesting topic about advanced waves. After reading all these, this author believes the advanced wave really exists. After this, this author proved that the mutual energy theorem is an energy theorem [1] by proving it is a sub-theorem of Poynting theorem [14]. We know Poynting theorem is an energy theorem, hence, the mutual energy theorem is also an energy theorem. After these, this author found that the Poynting theorem conflicts with the energy conservation for N charges in the empty space scenario. In order to solve this conflict, this author introduced the self-energy principle [1], which says corresponding to the retarded wave and advanced wave of the electromagnetic field, there are two time-reversal waves which cancel the self-energy flow of the retarded and advanced wave. Hence, the retarded wave and advanced wave do not carry any energy, they are probability waves instead of energy waves.

Maxwell equations together with the self-energy principle can derive the mutual energy principle [1]. The mutual energy principle can also derive the Maxwell equations. However, the solutions of the mutual energy principle, are different with Maxwell equations. The Maxwell equations have two solutions, the retarded solution and the advanced solution, which exist independently. The solutions of the mutual energy principle are also the retarded solution and the advanced solution, but these two solutions exist together and must be synchronized together. The synchronized retarded wave and advanced wave can build the mutual energy flow which satisfies the mutual energy flow theorem. The mutual energy flow theorem is an extension of the mutual energy theorem. All these theories can be found in [1]. Details can also be found in [15] [16] [17].

1.3. Normalized Mutual Energy Flow

It should be clear in the mutual energy theorem, mutual energy flow theorem, mutual energy principle, the word "mutual" is only because the historical reason, this word can be dropped. It can be called as "energy theorem", "energy flow theorem", "energy principle". This is because the self-energy principle has told us the self-energy flow does not carry and transfer any energy, the energy is only carried and transferred by the mutual energy flow. Hence, the mutual energy flow is the exclusive energy flow. This energy flow has similar property of the photon. For example, the photon is substance that can transfer energy from a point (emitter) to another point (absorber). The photon energy that goes through each surface between the emitter and the absorber should be equal, which is the energy of the photon. This property of photon is just described by the mutual energy flow theorem. Hence, this author claimed that the photon is the mutual energy flow.

However, the mutual energy flow calculated by the classical electromagnetic field theory is decreased when the distance between the emitter and the absorber is increased. We know photon energy does not decrease with this distance and is a constant. Considering this, in this article, this author would like to normalize the mutual energy flow. After the normalization, the normalized mutual energy flow has exactly same properties as the photon.

It should be clear that, if the theory of mutual energy flow theorem, mutual energy principle, can still be put inside the frame of the Maxwell theory, the normalized mutual energy flow belongs to pure quantum effect and cannot be put inside the frame of Maxwell theory. However, this quantum effect has not been described by the current quantum mechanics theory. The preprint of this article can be found in ref [19].

2. Mutual Energy Theorem and Mutual Energy Flow Theorem

2.1. Mutual Energy Theorem

The Welch's reciprocity theorem can be written as,

$$\left(\xi_1(t),\tau_2(t)\right)=-\left(\tau_1(t),\xi_2(t)\right) \tag{1}$$

where t is time and,

$$\xi_i \equiv \left[\boldsymbol{E}_i,\boldsymbol{H}_i\right],\quad i=1,2$$

$$\tau_i \equiv \left[\boldsymbol{J}_i,\boldsymbol{K}_i\right],\quad i=1,2$$

$\boldsymbol{E}_i$ is electric field, $\boldsymbol{H}_i$ is magnetic field. $\boldsymbol{J}_i$ is electric current. $\boldsymbol{K}_i$ is magnetic current, which is 0 here. The two inner products are defined as,

$$\left(\xi_1(t),\tau_2(t)\right)\equiv\int_{t=0}^{\infty}\mathrm{d}t\iiint_{V_2}\boldsymbol{E}_1(t)\cdot\boldsymbol{J}_2(t)\mathrm{d}V \tag{2}$$

$$\left(\tau_1(t),\xi_2(t)\right)\equiv\int_{t=0}^{\infty}\mathrm{d}t\iiint_{V_1}\boldsymbol{J}_1(t)\cdot\boldsymbol{E}_2(t)\mathrm{d}V \tag{3}$$

$\boldsymbol{J}_1$ is inside volume V_1. $\boldsymbol{J}_2$ is inside V_2. The above formula. It should be clear that the above theorem can be easily extended to de Hoop's cross-correlation reciprocity theorem,

$$\left(\xi_1(t+\tau),\tau_2(t)\right)=-\left(\tau_1(t+\tau),\xi_2(t)\right) \tag{4}$$

The Fourier transform of the de Hoop's reciprocity theorem is the mutual energy theorem, the mutual energy theorem is in the Fourier domain, which still can be expressed as Equation (1). But with the meaning of inner product is in the Fourier transform domain, ω is angle frequency.

$$\left(\xi_1(\omega),\tau_2(\omega)\right)=-\left(\tau_1(\omega),\xi_2(\omega)\right) \tag{5}$$

$$\left(\tau_1(\omega),\xi_2(\omega)\right)\equiv\iiint_{V_1}\boldsymbol{J}_1(\omega)\cdot\boldsymbol{E}_2^*(\omega)\mathrm{d}V \tag{6}$$

$$\left(\xi_1(\omega),\tau_2(\omega)\right)\equiv\iiint_{V_2}\boldsymbol{E}_1(\omega)\cdot\boldsymbol{J}_2^*(\omega)\mathrm{d}V \tag{7}$$

The above is Rumsey reciprocity theorem and also the mutual energy of this author. The above 4 theorems (1, 4, 5) can be seen as same theorem in two different domain, time-domain and Fourier domain. It is clear that these theorems are reciprocity theorems, but this author claimed that it is also an energy theorem [10] and hence, called it as the mutual energy theorem. In the above theorem, this author usually assumed τ_1 is the source, its field ξ_1 is the retarded field, τ_2 is the sink, its field ξ_2 is the advanced field. This theorem tells us that the advanced wave ξ_2 sacked energy from the source τ_1 which is equal to the energy that the retarded field ξ_1 offers to the sink τ_2. In the formulas (1, 4, 5) there is a negative sign, that is because τ_1 is source, the source offers energy, hence, has a negative sign. Here the tradition is that the consumed energy is positive, hence, the energy offered by the source is negative. The right of the formula (1, 4, 5) is the energy offered by the source τ_1. The left of the formula (1, 3, 5) is the consumed energy by the current of the sink τ_2. The theorem tells us these two energies are equal.

2.2. Mutual Energy Flow Theorem

The above theorem has been extended as mutual energy flow theorem [1],

$$\left(\xi_1(t),\tau_2(t)\right)=\left(\xi_1(t),\xi_2(t)\right)=-\left(\tau_1(t),\xi_2(t)\right) \tag{8}$$

The mutual energy flow is defined as

$$\left(\xi_1(t),\xi_2(t)\right)\equiv\int_{t=0}^{\infty}\oiint_{\Gamma}\left(\boldsymbol{E}_1(t)\times\boldsymbol{H}_2(t)+\boldsymbol{E}_2(t)\times\boldsymbol{H}_1(t)\right)\cdot\hat{n}\mathrm{d}\Gamma\mathrm{d}t \tag{9}$$

Γ is the arbitrary close surface or open surface which has extended to infinity between the source τ_1 and the sink τ_2, see Figure 1.

The mutual energy flow theorem tells us, that the energy from the source to the sink is transferred through the mutual energy flow. The mutual energy flow is also a good inner product of the two fields, $\xi_1(t),\xi_2(t)$. The mutual energy flow theorem can also be converted to Fourier domain,

$$\left(\xi_1(\omega),\tau_2(\omega)\right)=\left(\xi_1(\omega),\xi_2(\omega)\right)=-\left(\tau_1(\omega),\xi_2(\omega)\right) \tag{10}$$

$$\left(\xi_1(\omega),\xi_2(\omega)\right)\equiv\oiint_{\Gamma}\left(\boldsymbol{E}_1(\omega)\times\boldsymbol{H}_2^*(\omega)+\boldsymbol{E}_2^*(\omega)\times\boldsymbol{H}_1(\omega)\right)\cdot\hat{n}\mathrm{d}\Gamma \tag{11}$$

The mutual energy flow theorem can be seen by Figure 1. In the Figure **a** is the source or emitter and **b** is the sink or absorber. The emitter can be seen as an atom with a transmitting antenna inside it. The absorber can be seen as an atom with a receiving antenna inside it. The source sends the retarded wave. The sink sends the advanced wave. The mutual energy flow is built by the retarded wave and advanced wave. The shape of the mutual energy flow is the overlap of retarded wave and the advanced wave which are thin in the two ends and thick in the middle. The mutual energy goes through any surfaces between **a** and **b** are equal. This is the mutual energy flow theorem. The mutual energy flow theorem further guarantees that the mutual energy theorem is an energy theorem.

3. Normalized Mutual Energy Flow

3.1. Normalized Mutual Energy Flow

This author has claimed that the photon is nothing else, but the mutual energy flow [1]. However, recently this author also noticed that the mutual energy flow calculated according to classical electromagnetic field theory decreases when the distance between the source and sink increases. This is clear, since the radiation fields decreases

$$\xi_1(\boldsymbol{x}_2)\propto\frac{1}{r} \tag{12}$$

$$\xi_2(\boldsymbol{x}_1)\propto\frac{1}{r} \tag{13}$$

$$r\equiv\left\|\boldsymbol{x}_2-\boldsymbol{x}_1\right\| \tag{14}$$

$\boldsymbol{x}_1$ is the position of source. $\boldsymbol{x}_2$ is the position of sink. In other side we know the mutual energy flow should be the photon which has fixed energy $E=\hbar\omega$, where $\hbar$ is reduced Plank constant, ω is the angle frequency of the light. Hence, we require that,

$$\left(\xi_1(t),\xi_2(t)\right)=\hbar\omega \tag{15}$$

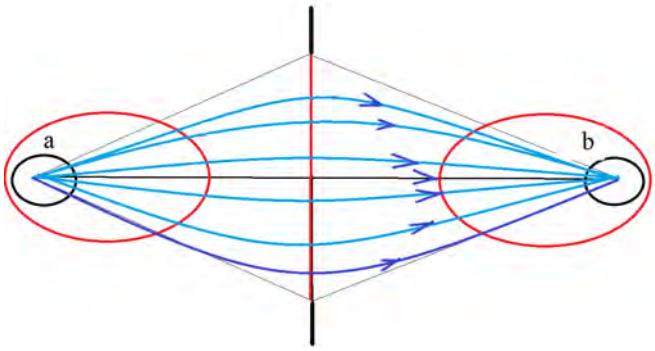


Figure 1. The mutual energy flow, which is started from source "a" and ended at sink "b". Source "a" is an emitter or transmitting antenna. Sink "b" is absorber or a receiving antenna. Photon is the mutual energy flow. According to the mutual energy flow theorem, the mutual energy go through any of the surfaces between source "a" and sink "b" are equal. The surfaces can be closed surface for example the sphere surface surrounding to the source "a" or the sink "b", or the infinite plane between "a" and "b".

In order to satisfy the above formula, we need to adjust the value of fields. We need to increase the fields according to the distance $r \equiv \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$, that is,

$$kr\left(\xi_1(t), \tau_2(t)\right) = kr\left(\xi_1(t), \xi_2(t)\right) = -kr\left(\tau_1(t), \xi_2(t)\right) = \hbar\omega \tag{16}$$

k is a constant which is not related the distance r. The above can be re-written as,

$$\left(\xi_1^p(t), \tau_2^p(t)\right) = \left(\xi_1^p(t), \xi_2^p(t)\right) = -\left(\tau_1^p(t), \xi_2^p(t)\right) = \hbar\omega \tag{17}$$

where,

$$\xi_i^p \equiv \sqrt{kr}\,\xi_i(t), \ i = 1,2 \tag{18}$$

$$\tau_i^p \equiv \sqrt{kr}\,\tau_i(t), \ i = 1,2 \tag{19}$$

ξ_i^p is the electromagnetic fields of the photon, τ_i^p is the source of the photon. The field of photon is much stronger than the mutual energy flow calculated according to the classical electromagnetic fields, especially when the distance $r = \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$ is very large. We can think that the energy is carried by the mutual energy flow of the classical electromagnetic field which is very small. However, this mutual energy flow builds an energy channel from emitter to the absorber. The energy channel can transfer the energy from emitter (source) to the absorber (sink). This energy channel can be applied repeatedly until the energy reach to the energy of one photon. That is the reason the photon field is stronger than the classical electromagnetic fields.

3.2. Normalization the Wave Function in Quantum Mechanics

In quantum mechanics there is also the normalization of the wave function. The square of amplitude of the wave function is normalized to 1. This 1 is probability. The square of amplitude of wave function is normalized to 1 is similar to the normalized mutual energy flow to the energy of one photon. The difference between the two ways are only a constant.

Because this author believes the wave is an energy-based wave instead of probability-based wave, the photon was normalized to the energy $\hbar\omega$. It should be clear that after the normalization, the field of photon ξ_i^p is quite different with the classical electromagnetic field theory ξ_i. The photon field is normalized electromagnetic fields.

It should be clear, that in the current example the retarded field $\xi_1(\boldsymbol{x})=[\boldsymbol{E}_1,\boldsymbol{H}_1]$ is sent from the source **a** and hence, decreases according to $\|\boldsymbol{x}-\boldsymbol{x}_1\|$. The advanced field $\xi_2(\boldsymbol{x})=[\boldsymbol{E}_2,\boldsymbol{H}_2]$ is sent from the sink **b** and hence, decreases according to $\|\boldsymbol{x}-\boldsymbol{x}_2\|$. Even $\xi_1(\boldsymbol{x})$ and $\xi_2(\boldsymbol{x})$ decrease with the distance $\|\boldsymbol{x}-\boldsymbol{x}_1\|$, $\|\boldsymbol{x}-\boldsymbol{x}_2\|$, the mutual energy flow $(\xi_2(\boldsymbol{x}),\xi_1(\boldsymbol{x}))$ is not changed at any surfaces see **Figure 1**.

There is also another possibility that the advanced wave become the guidance wave of the retarded wave and the retarded wave become the guidance wave of the advanced wave. Because of the effect of guidance waves, the retarded wave and the advanced wave both do not decrease with the distance $\|\boldsymbol{x}-\boldsymbol{x}_1\|$ and $\|\boldsymbol{x}-\boldsymbol{x}_2\|$. However, inside the photon we cannot measure the field $[\boldsymbol{E}_1,\boldsymbol{H}_1]$ and $[\boldsymbol{E}_2,\boldsymbol{H}_2]$. This scenario has the same mutual energy flow as before. For calculating the mutual energy flow we still can use Equation (9).

4. The Macroscopic Electromagnetic Field

Here macroscopic electromagnetic field means the field produced by many photons together. This section we prove that the macroscopic electromagnetic field satisfy Maxwell equations of the retarded wave. We need a condition: all absorbers are uniformly distributed on the big sphere with its radius as infinite.

4.1. Poynting Theorem Is Equivalent to Maxwell Equations

It is known that the Maxwell equations,

$$\nabla\times\boldsymbol{E}=-\partial_t\boldsymbol{B} \tag{20}$$

$$\nabla\times\boldsymbol{H}=\boldsymbol{J}+\partial_t\boldsymbol{D} \tag{21}$$

$$\boldsymbol{B}=\mu_0\boldsymbol{H},\ \ \boldsymbol{D}=\epsilon_0\boldsymbol{E} \tag{22}$$

(where $\partial_t=\dfrac{\partial}{\partial t}$) is equivalent to the corresponding to Poynting theorem [18],

$$-\oiint_\Gamma \boldsymbol{E}\times\boldsymbol{H}=\iiint_V(\boldsymbol{E}\cdot\boldsymbol{J}+\boldsymbol{E}\cdot\partial_t\boldsymbol{D}+\boldsymbol{H}\cdot\partial_t\boldsymbol{B})\mathrm{d}V \tag{23}$$

If we can prove the field of many photons satisfy Poynting theorem above we have proved that the macroscopic electromagnetic field satisfy Maxwell equations.

4.2. The Self-Energy Principle and the Mutual Energy Principle

Assume in the space there are N charges. For the electromagnetic field of each charge, it satisfies Maxwell equations and hence, the Poynting theorem,

$$-\oiint_{\Gamma} \boldsymbol{E}_i \times \boldsymbol{H}_i \cdot \hat{n} \mathrm{d}\Gamma = \iiint_V \left(\boldsymbol{E}_i \cdot \boldsymbol{J}_i + \boldsymbol{E}_i \cdot \partial_t \boldsymbol{D}_i + \boldsymbol{H}_i \cdot \partial_t \boldsymbol{B}_i \right) \mathrm{d}V \tag{24}$$

Substitute the superposition principle

$$\boldsymbol{E} = \sum_{i=1}^{N} \boldsymbol{E}_i, \quad \boldsymbol{H} = \sum_{i=1}^{N} \boldsymbol{H}_i \tag{25}$$

to the Poynting theorem Equation (23) we obtain

$$\begin{aligned} &-\sum_{j=1}^{N} \sum_{i=1}^{N} \oiint_{\Gamma} \boldsymbol{E}_j \times \boldsymbol{H}_i \cdot \hat{n} \mathrm{d}\Gamma \\ &= \sum_{j=1}^{N} \sum_{i=1}^{N} \iiint_V \left(\boldsymbol{E}_j \cdot \boldsymbol{J}_i + \boldsymbol{E}_j \cdot \partial_t \boldsymbol{D}_i + \boldsymbol{H}_j \cdot \partial_t \boldsymbol{B}_i \right) \mathrm{d}V \end{aligned} \tag{26}$$

Equation (24) is the energy of self-energy flow, we have introduced the self-energy principle [1] which tell us that self-energy flows do not carry any energy. The self-energy flow terms (which have $i = j$) have been all canceled by corresponding terms of the self-energy flows of the time-reversal waves. Hence, we can take away all self-energy flow terms from Equation (26), we obtain,

$$\begin{aligned} &-\sum_{j=1}^{N} \sum_{i=1, i \neq j}^{N} \oiint_{\Gamma} \boldsymbol{E}_j \times \boldsymbol{H}_i \cdot \hat{n} \mathrm{d}\Gamma \\ &= \sum_{j=1}^{N} \sum_{i=1, i \neq j}^{N} \iiint_V \left(\boldsymbol{E}_j \cdot \boldsymbol{J}_i + \boldsymbol{E}_j \cdot \partial_t \boldsymbol{D}_i + \boldsymbol{H}_j \cdot \partial_t \boldsymbol{B}_i \right) \mathrm{d}V \end{aligned} \tag{27}$$

Equation (27) is the mutual energy formula, we have promoted it as the mutual energy principle, and claim it is real physical equation for photon [1]. From mutual energy principle, we can derive the Maxwell equations Equation (20, 21). However, the derived Maxwell equations must belong to the two groups, one is the retarded wave and another is the advanced wave. The retarded wave must synchronize with the advanced wave. The synchronized retarded wave and advanced wave can produce the mutual energy flow which satisfy the mutual energy flow theorem. Hence, the mutual energy principle is the foundation of the mutual energy flow theorem which describes the photon. In the follow section we will prove the macroscopic field *i.e.* the field of many photons that satisfy Maxwell equations by using the mutual energy principle. In this way, we have proved that the macroscopic wave can be built by thousands of photons together and proves Einstein's guesses.

4.3. The Macroscopic Field of the Electromagnetic Field

Equation (27) can be re-written as,

$$\begin{aligned} &-\sum_{j=1}^{N} \sum_{i=1}^{i<j} \oiint_{\Gamma} \left(\boldsymbol{E}_j \times \boldsymbol{H}_i + \boldsymbol{E}_i \times \boldsymbol{H}_j \right) \cdot \hat{n} \mathrm{d}\Gamma \\ &= \sum_{j=1}^{N} \sum_{i=1}^{i<j} \iiint_V \left(\boldsymbol{E}_j \cdot \boldsymbol{J}_i + \boldsymbol{E}_i \cdot \boldsymbol{J}_j + \boldsymbol{E}_j \cdot \partial_t \boldsymbol{D}_i + \boldsymbol{E}_i \cdot \partial_t \boldsymbol{D}_j + \boldsymbol{H}_j \cdot \partial_t \boldsymbol{B}_i + \boldsymbol{H}_i \cdot \partial_t \boldsymbol{B}_j \right) \mathrm{d}V \end{aligned} \tag{28}$$

Considering among these charges there are L emitters and M absorbers. The retarded field of the emitter can only paired to an advanced field from absorbers, hence Equation (28) can be re-written as,

$$\begin{aligned} &-\sum_{j=1}^{L} \sum_{i=1}^{M} \oiint_{\Gamma} \left(\boldsymbol{E}_j \times \boldsymbol{H}_i + \boldsymbol{E}_i \times \boldsymbol{H}_j \right) \cdot \hat{n} \mathrm{d}\Gamma \\ &= \sum_{j=1}^{L} \sum_{i=1}^{M} \iiint_V \left(\boldsymbol{E}_j \cdot \boldsymbol{J}_i + \boldsymbol{E}_i \cdot \boldsymbol{J}_j + \boldsymbol{E}_j \cdot \partial_t \boldsymbol{D}_i + \boldsymbol{E}_i \cdot \partial_t \boldsymbol{D}_j + \boldsymbol{H}_j \cdot \partial_t \boldsymbol{B}_i + \boldsymbol{H}_i \cdot \partial_t \boldsymbol{B}_j \right) \mathrm{d}V \end{aligned} \tag{29}$$

We consider a special situation, the source is at the origin point, there are infinite more absorbers are at the big sphere with its radius infinite large. Assume $L = 1$, there is only one source (Emitter) which radiate the retarded wave. All other M charges are absorbers which are on the big sphere. In this special situation the above formula can be written as,

$$\begin{aligned}&-\sum_{i=1}^{M}\oiint_{\Gamma}\left(\boldsymbol{E}_0\times\boldsymbol{H}_i+\boldsymbol{E}_i\times\boldsymbol{H}_0\right)\cdot\hat{n}\mathrm{d}\Gamma\\&=\sum_{i=1}^{M}\iiint_{V}\left(\boldsymbol{E}_0\cdot\boldsymbol{J}_i+\boldsymbol{E}_i\cdot\boldsymbol{J}_0+\boldsymbol{E}_0\cdot\partial_t\boldsymbol{D}_i+\boldsymbol{E}_i\cdot\partial_t\boldsymbol{D}_0+\boldsymbol{H}_0\cdot\partial_t\boldsymbol{B}_i+\boldsymbol{H}_i\cdot\partial_t\boldsymbol{B}_0\right)\mathrm{d}V\end{aligned}\tag{30}$$

In the above, the retarded wave has the subscribe "0". We have considered the advanced wave cannot build the mutual energy with advanced waves; hence, all absorber must build mutual energy with the only source. The above can be written as,

$$\begin{aligned}&-\oiint_{\Gamma}\left(\boldsymbol{E}_0\times\boldsymbol{H}_a+\boldsymbol{E}_a\times\boldsymbol{H}_0\right)\\&=\iiint_{V}\left(\boldsymbol{E}_0\cdot\boldsymbol{J}_a+\boldsymbol{E}_a\cdot\boldsymbol{J}_0+\boldsymbol{E}_0\cdot\partial_t\boldsymbol{D}_a+\boldsymbol{E}_a\cdot\partial_t\boldsymbol{D}_0+\boldsymbol{H}_0\cdot\partial_t\boldsymbol{B}_a+\boldsymbol{H}_a\cdot\partial_t\boldsymbol{B}_0\right)\mathrm{d}V\end{aligned}\tag{31}$$

where $\boldsymbol{E}_a=\sum_{i=1}^{M}\boldsymbol{E}_i$, $\boldsymbol{H}_a=\sum_{i=1}^{M}\boldsymbol{H}_i$ are superposed advanced wave. In the above formula, we can take the volume only includes the source. In this situation, the above formula become,

$$\begin{aligned}&-\oiint_{\Gamma}\left(\boldsymbol{E}_0\times\boldsymbol{H}_a+\boldsymbol{E}_a\times\boldsymbol{H}_0\right)\cdot\hat{n}\mathrm{d}\Gamma\\&=\iiint_{V}\left(\boldsymbol{E}_a\cdot\boldsymbol{J}_0+\boldsymbol{E}_0\cdot\partial_t\boldsymbol{D}_a+\boldsymbol{E}_a\cdot\partial_t\boldsymbol{D}_0+\boldsymbol{H}_0\cdot\partial_t\boldsymbol{B}_a+\boldsymbol{H}_a\cdot\partial_t\boldsymbol{B}_0\right)\mathrm{d}V\end{aligned}\tag{32}$$

$\boldsymbol{E}_0\cdot\boldsymbol{J}_a$ is disappear in the above formula, since J_a are at the big sphere which is at outside of V. In this case the absorbers uniformly distribute on the big sphere, the advanced wave will exactly same as the retarded field radiate from the source which is at the origin. Hence, the above formula can be written as,

$$\begin{aligned}&-\oiint_{\Gamma}\left(\boldsymbol{E}_0\times\boldsymbol{H}_0+\boldsymbol{E}_0\times\boldsymbol{H}_0\right)\cdot\hat{n}\mathrm{d}\Gamma\\&=\iiint_{V}\left(\boldsymbol{E}_0\cdot\boldsymbol{J}_0+\boldsymbol{E}_0\cdot\partial_t\boldsymbol{D}_0+\boldsymbol{E}_0\cdot\partial_t\boldsymbol{D}_0+\boldsymbol{H}_0\cdot\partial_t\boldsymbol{B}_0+\boldsymbol{H}_0\cdot\partial_t\boldsymbol{B}_0\right)\mathrm{d}V\end{aligned}\tag{33}$$

or

$$-2\oiint_{\Gamma}\left(\boldsymbol{E}_0\times\boldsymbol{H}_0\right)\cdot\hat{n}\mathrm{d}\Gamma=\iiint_{V}\left(\boldsymbol{E}_0\cdot\boldsymbol{J}_0+2\left(\boldsymbol{E}_0\cdot\partial_t\boldsymbol{D}_0+\boldsymbol{H}_0\cdot\partial_t\boldsymbol{B}_0\right)\right)\mathrm{d}V\tag{34}$$

Assume the total electromagnetic field together of retarded field and the advanced field are

$$\boldsymbol{E}=\boldsymbol{E}_0+\boldsymbol{E}_a,\quad \boldsymbol{H}=\boldsymbol{H}_0+\boldsymbol{H}_a$$

We have

$$\boldsymbol{E}=\boldsymbol{E}_0+\boldsymbol{E}_0=2\boldsymbol{E}_0,\quad \boldsymbol{H}=\boldsymbol{H}_0+\boldsymbol{H}_0=2\boldsymbol{H}_0\tag{35}$$

Substitute it to Equation (34), we have,

$$-\oiint_{\Gamma}\left(\boldsymbol{E}\times\boldsymbol{H}\right)\cdot\hat{n}\mathrm{d}\Gamma=\iiint_{V}\left(\boldsymbol{E}\cdot\boldsymbol{J}_0+\boldsymbol{E}\cdot\partial_t\boldsymbol{D}+\boldsymbol{H}\cdot\partial_t\boldsymbol{B}\right)\mathrm{d}V\tag{36}$$

This is the Poynting theorem of the total electromagnetic fields $\boldsymbol{E},\boldsymbol{H}$. $\boldsymbol{J}_0$ is the current intensity of the source. Please notice we have assumed that the absorbers are uniformly distributed on the big sphere. In this situation, the

total electromagnetic field which are sum of the retarded wave from the emitter and the advanced wave from the absorber satisfy the Poynting theorem. If this electromagnetic field satisfies Poynting theorem, it also satisfies the Maxwell equations. It seems that all the electromagnetic fields are started from the source. Hence, we have proved that the electromagnetic fields of a point source together with the fields of the absorbers at big sphere with infinite radius satisfy Maxwell equations. This field looks like the retarded field. Here it only looks like the retarded field. The fields $\boldsymbol{E},\boldsymbol{H}$ actually are the summation of the retarded field radiate from the source and advanced wave radiate from the absorbers.

If the macroscopic electromagnetic fields of a point source with the field of the absorber satisfies Maxwell equations, we can extend this result to many point sources inside a region. It should be clear, here the total electromagnetic field are contributions of the retarded wave and the advanced wave, but it looks like as all the field are send as retarded wave from the emitter. Hence, for macroscopic wave, it still can be seen as retarded. Since the mutual energy flow are the energy flow of the photon, we have proved that the macroscopic wave can be built by many photons. This macroscopic wave can be seen as retarded wave even it is built from the mutual energy flows which are in turn built from retarded wave and advanced wave.

4.4. The Effect of the Normalized Mutual Energy Flow

In the above we have actually proved the total electromagnetic fields satisfy Maxwell equations. The total electromagnetic field has two parts one is the retarded field from the source, the other are advanced wave from the uniformly distributed absorbers on a big sphere. Normalized mutual energy flow is the photon energy, which is much stronger than the mutual energy flows calculated through classical electromagnetic field theory without the normalization. The number of the normalized mutual energy flows are also much less the mutual energy flows calculated from classical electromagnetic field theory. If the number of normalized mutual energy flows are enough more, they will have same effect as the mutual energy flows calculated through the classical electromagnetic field theory (which has much smaller values, but much larger in numbers).

If the absorbers could be seen as uniformly distribute at big sphere, the advanced waves of the total absorbers will have the same value as the retarded wave. Hence, the above proof in subsection 4.3 is also effective for the case of normalized electromagnetic fields, which is the field of photons.

4.5. The Structure of the Macroscopic Electromagnetic Fields

According to the discussion of this article, the structure of the electromagnetic fields can be shown in **Figure 2**. From these, we can see the macroscopic electromagnetic fields have a complicated structure. The macroscopic electromagnetic fields are built by infinite mutual energy flows. The effect of

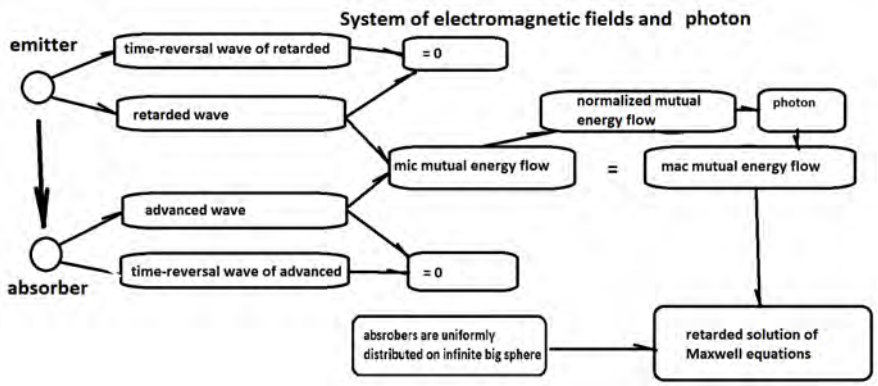


Figure 2. The structure of the electromagnetic field. In this figure the energy is sent from the emitter to the absorbers. Assume that the absorbers are uniformly distribute at big sphere with infinite radius. The emitter sends retarded wave and corresponding time-reversal wave. These two energy flows cancel. The absorber sends advanced wave and corresponding time-reversal wave. These two energy flows cancel. The retarded wave and the advanced wave interacted and produce the microscopic mutual energy flow. Normalized mutual microscopic energy flows are photons. Many photons have the same effect with many microscopic mutual energy flows. Macroscopic mutual energy flows have the same values as microscopic mutual energy flow. Macroscopic mutual energy flows can build the retarded solution of the Maxwell equations in case the absorber can be seen as uniformly distribute on the big sphere with infinite radius.

the mutual energy flow is same as infinite photons. It should be clear that the mutual energy flows are much smaller than the photon. However, when both mutual energy flows and photons have very big number, they are equivalent. In order to build the macroscopic wave to satisfy Maxwell equations, the absorbers have to be uniformly distributed on the big sphere with an infinite radius. Photons are the normalized mutual energy flows. After the normalization, the mutual energy flow has the energy of $E = \hbar\omega$. The mutual energy flows are built by the retarded wave sent from the emitter and the advanced wave sent from the absorber. The emitter also radiates the time-reversal wave which cancel the energy flow of the retarded wave. The absorber also radiates the time-reversal wave which cancel the energy flow of the advanced wave.

This figure tells us how macroscopic wave of electromagnetic field is built by photons which are normalized mutual energy flows. The mutual energy flow is the interaction of the retarded wave radiate from the source and the advanced wave radiate from the sink. The self-energy flow of the retarded wave and the advanced wave has been canceled by the corresponding time-reversal waves. Hence the waves, retarded and advanced wave, can be seen as probability wave, because they do not carry energy, if we consider the effect of the time-reversal waves. This means the microscopic waves are only look like probability wave, they are also real physic waves.

5. In Case the Absorber Cannot Be Seen as Uniformly Distributed

In order to test the above theory, we noticed that the above theory is only satis-

fied in case the absorbers can be seen as uniformly distributed in big sphere with infinite radius. In case this condition cannot be satisfied, the result should be not obtained. That means the macroscopic wave of many photons should not satisfy Maxwell equations. This is true. For example, we assume there is a transmitting antenna radiate the wireless wave. We assume the environment can be seen as uniformly distributed absorbers. Hence, according to the above theory, the electromagnetic fields includes the retarded wave sent by the transmitting antenna and the advanced wave sent by the uniformly distributed absorbers together can be seen as retarded waves sent out from the transmitting antenna. Now we put a receiving antenna inside the above space. It is clear that when the receiving antenna is put in the system, closing to the receiving antenna, the absorbers cannot be seen as uniformly distributes on the big sphere. This means that closing to the receiving antenna, the electromagnetic field should not satisfy the Maxwell equations of the retarded wave. This is true. We know that for the receiving antenna, there are a concept of effective scattering section area. If effective scattering section area equals to the section area of the antenna, it tells us the field close to the antenna exactly satisfies the Poynting theorem. If the effective scattering section area larger than the section area of the receiving antenna, that means the Poynting theorem has not been exactly satisfied. We know from electronic engineering, for the receiving wire antenna, the effective scattering section area can be a hundred times larger than its section area. This means in this case the Poynting theorem cannot be satisfied. Since the Poynting theorem is equivalent to the two Maxwell equations Equation (20, 21), that means the Maxwell equations are also not satisfied.

In electronic engineering, the directivity dialog of receiving antenna cannot be calculated according to the theory of the retarded wave of Maxwell equations. Through the experiment we find the directivity dialog of the receiving antenna is same as the antenna used as the transmitting antenna. Hence engineers calculate the directivity dialog of the receiving antenna just using the same method as it is a transmitting antenna. This way it needs a proof that the directivity dialog of the receiving antenna is same as that of the transmitting antenna. Normally Lorentz reciprocity theorem is used for this task. It should be clear that the mutual energy flow theorem can also do this job. Since, in the directivity dialog is absolute value, and the phase information is omitted, and hence have the same value for the mutual energy theorem and Lorenz reciprocity theorem. Hence, the directivity dialog calculation for the receiving antenna is also because of the mutual energy theorem not Poynting theorem and also not Maxwell equations.

6. Conclusions

It is found that photon can be described by the normalized mutual energy flow. The mutual energy flow is the interaction of the retarded wave radiated from the source and the advanced wave radiated from the absorber. The normalization process is a process of physics which cannot be put in the frame of Maxwell theory. It is a pure quantum mechanics effect.

After the normalization of the mutual energy flow, the mutual energy flow can be a good description of photon. The structure of the macroscopic electromagnetic field becomes clear; the macroscopic electromagnetic field is built by many photons. The effect of many photons is same as many mutual energy flows without the normalization. The mutual energy flow is built by the retarded wave and the advanced wave. The self-energy flow of the retarded wave and the advanced wave have canceled by the corresponding energy flow of the time-reversal waves of the retarded wave and the advanced wave.

With the theory of mutual energy flow and this normalization, the concept of collapse of the wave is not required. The wave and particle duality paradox have been solved.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Zhao, S.R. (2017) *American Journal of Modern Physics and Application*, **4**, 12-23.

[2] Wheeler, J.A. and Feynman, R.P. (1945) *Reviews of Modern Physics*, **17**, 157. https://doi.org/10.1103/RevModPhys.17.157

[3] Wheeler, J.A. and Feynman, R.P. (1949) *Reviews of Modern Physics*, **21**, 425. https://doi.org/10.1103/RevModPhys.21.425

[4] Schwarzschild, K. (1903) *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, **128**, 132-141.

[5] Tetrode, H. (1922) *Zeitschrift für Physik*, **10**, 137. https://doi.org/10.1007/BF01332555

[6] Fokker, A.D. (1929) *Zeitschrift für Physik*, **58**, 386. https://doi.org/10.1007/BF01340389

[7] Cramer, J. (1986) *Reviews of Modern Physics*, **58**, 647-688. https://doi.org/10.1103/RevModPhys.58.647

[8] Welch, W.J. (1960) *IRE Transactions on Antennas and Propagation*, **8**, 68-73. https://doi.org/10.1109/TAP.1960.1144806

[9] Rumsey, V.H. (1963) *IEEE Transactions on Antennas and Propagation*, **11**, 73-86. https://doi.org/10.1109/TAP.1963.1137982

[10] Zhao, S.R. (1987) *Acta Electronica Sinica*, **15**, 88-93.

[11] de Hoop, A.T. (1987) *Radio Science*, **22**, 1171-1178. https://doi.org/10.1029/RS022i007p01171

[12] Zhao, S.R. (1989) *Journal of Electronics*, **11**, 204-208.

[13] Zhao, S.R. (1989) *Journal of Electronics*, **11**, 73-77.

[14] Stephenson, L.M. (2000) *Physics Essays*, **13**, 138.

[15] Zhao, S.R. (2020) Electromagnetic Field, Photon and Quantum with Advanced Wave and Time-Reversal Wave Based on Mutual Energy Principle and Self-Energy Principle (Volume I). https://hal.archives-ouvertes.fr/hal-02512141

[16] Zhao, S.R. (2020) Electromagnetic Field, Photon and Quantum with Advanced

Wave and Time-Reversal Wave Based on Mutual Energy Principle and Self-Energy Principle (Volume II). https://hal.archives-ouvertes.fr/hal-02512153

[17] Zhao, S.R. (2020) Electromagnetic Field, Photon and Quantum with Advanced Wave and Time-Reversal Wave Based on Mutual Energy Principle and Self-Energy Principle (Volume III).

[18] Poynting, J.H. (1884) *Philosophical Transactions of the Royal Society of London*, **175**, 343-361. https://doi.org/10.1098/rstl.1884.0016

[19] Zhao, S.R. (2019) Photon Is Interpreted by the Particleization/Normalization of the Mutual Energy Flow of the Electromagnetic Fields. https://vixra.org/abs/1910.0079

Journal of Modern Physics, 2020, 11, 683-702
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Kolmogorov's Probability Spaces for "Entangled" Data-Subsets of EPRB Experiments: No Violation of Einstein's Separation Principle

Karl Hess

Center for Advanced Study, University of Illinois, Urbana, Illinois, USA
Email: karlfhess@gmail.com

How to cite this paper: Hess, K. (2020) Kolmogorov's Probability Spaces for "Entangled" Data-Subsets of EPRB Experiments: No Violation of Einstein's Separation Principle. *Journal of Modern Physics*, **11**, 683-702.
https://doi.org/10.4236/jmp.2020.115044

Received: April 15, 2020
Accepted: May 6, 2020
Published: May 9, 2020

Abstract

It is demonstrated that the use of Kolmogorov's probability theory to describe results of quantum probability for EPRB (Einstein-Podolsky-Rosen-Bohm) experiments requires extreme care when different subsets of measurement outcomes are considered. J. S. Bell and his followers have committed critical inaccuracies related to spin-gauge and probability measures of such subsets, because they use exclusively a single probability space for all data sets and sub-sets of data. It is also shown that Bell and followers use far too stringent epistemological requirements for the consequences of space-like separation. Their requirements reach way beyond Einstein's separation principle and cannot be met by the major existing physical theories including relativity and even classical mechanics. For example, the independent free will does not empower the experimenters to choose multiple independent spin-gauges in the two EPRB wings. It is demonstrated that the suggestion of instantaneous influences at a distance (supposedly "derived" from experiments with entangled quantum entities) is a consequence of said inaccuracies and takes back rank as soon as the Kolmogorov probability measures are related to a consistent global spin-gauge and permitted to be different for different data subsets: Using statistical interpretations and different probability spaces for certain subsets of outcomes instead of probability amplitudes related to single quantum entities, permits physical explanations without a violation of Einstein's separation principle.

Keywords

Bell's Theorem, Einstein's Separation Principle, EPRB Experiments

1. Introduction

The differences between classical probability and the modified probabilistic

concepts used in quantum mechanics, have been the topic of many discussions related to the foundations of quantum mechanics and are in a way the root of Feynman's well-known remark that no one understands quantum mechanics. To avoid ambiguities, I define Kolmogorov's (set-theoretic) probability framework as "classical" probability and the absolute square of Feynman's probability amplitudes as the quantum probability version.

The classical-quantum distinction of probabilistic concepts has appeared in clearest relief due to the work of J. S. Bell [1] and his inequalities that involve Einstein-Podolsky-Rosen (EPR) experiments [2] and corresponding measurements of entangled pairs. Wigner [3] presented a set-theoretic version of Bell's inequalities. The actually performed experiments (see e.g. [4]) are a variation of EPR as proposed by Bohm (EPRB). The experimental results contradict the Bell-Wigner inequalities, a contradiction that has led to the common belief that instantaneous influences at a distance are at work in experiments of entangled quantum "entities" (photons, electrons etc.) and that Einstein's separation principle derived from the speed c of light in vacuum is actually violated in EPRB experiments.

Any criticism of the Bell-Wigner inequalities is currently seen by a majority of physicists as nonsensical and comparable to the attempts to build a perpetuum mobile that contradicts energy conservation. It is the purpose of this paper to show that the criticism of the Bell-Wigner approach should rather be compared to the early days of the calculus and to Berkeley's criticism of Newton's fluxions and Leibniz's infinitesimals that were put to zero after a logical procedure that regarded them as definitely not-zero. Bishop Berkeley stated that such a method of reasoning would not be allowed in Divinity. It took about a century and the work of Cauchy, Weierstrass and other notables to repair the problems convincingly.

Instantaneous influences over space-like distances do have similar logical problems as the fluxions of Newton and are indeed considered instantaneous only when the outcomes of the influences are random and when it is absolutely impossible to transmit any information instantaneously. Any instantaneous transfer of information would contradict Einstein's relativity and no sane physicist believes in such a possibility. Therefore, in a variety of descriptions, "influences" are introduced that are "instantaneous" only if the instantaneity cannot directly be proven but only statistically inferred. That statistical inference is invariably based on the violation of the Bell-Wigner type of inequalities.

It will be shown in great detail below that the commonly used logic of applying Bell-Wigner-type inequalities to actual EPRB experiments and/or the results of quantum mechanics would also not be allowed in Divinity because of a variety of reasons that uncover serious inaccuracies of epistemological, physical and mathematical nature. In contrast to the case of calculus, a repair of these problems appears unlikely.

Avoiding the mentioned inaccuracies permits us to construct a model for EPRB experiments, which is based on a statistical relation of certain subsets of

measurement outcomes as opposed to interpretations regarding individual entangled pairs.

2. EPRB Experiments and Notation

I assume that the reader is reasonably familiar with EPRB experiments that deal with entangled (correlated) photon pairs. These pairs are sent into two directions or into two optical fibers. To be definite we assume that one of the entangled photons propagates perpendicular to the x, y plane that labels one face (perpendicular to the z-direction) of a cube-like Wollaston prism. The x-direction of this chosen coordinate system may also be called "horizontal" and the y-direction "vertical". The second photon of the entangled pair propagates perpendicular to the face of a second Wollaston prism and that face is labeled by the x', y' coordinate system, which is perpendicular to the z' direction and we may use for simplicity $z = z'$. The photons exit the Wollaston cubes along two different directions and are registered by detectors D_v^1 and D_h^1 in wing 1 of the experimental system and by detectors D_v^2 and D_h^2 in wing 2, respectively. The measurement system must also guarantee that the signals detected in the two different and spatially distant wings belong to entangled pairs. This is usually achieved by registering the measurement time and by correlating the signals by the usual space-time correlations of photons propagating with the speed of light. Furthermore, we can find an orientation of maximal correlation by fixing Wollaston prism 1 and rotating Wollaston prism 2 perpendicular to the propagation direction of the photons until we have a virtually perfect anti-correlation of the outcomes, *i.e.* when the horizontal detector clicks in wing 1, the vertical detector clicks in wing 2 or vice versa. In actual experiments this happens for about 99% of all photon pairs or even better, while in the theories that we discuss this must happen with probability 1. It is convenient to define the directions of maximal anti-correlation in wing 2 also as the x, y-directions.

As we will see later, the theoretical work of Bell deals with at least two different rotations of the Wollaston prisms for each wing and relates the corresponding detector registrations to a "horizontal" and "vertical" spin component for all the different orientations of the Wollaston prisms. Thus, we do not have a carefully defined and unique gauge for the spin measurements in Bell's work (in its most basic definition of "gauge" like that of the meter-measure in Paris) and we will see that this fact alone calls for extreme caution when formulating Bell's inequalities and using Einstein's hypothesis that elements of physical reality determine the spin outcomes and their correlations in both wings. It is certainly, in general logically inadmissible to denote different directions in a given wing by "horizontal" or "vertical". The gauge in the other wing is codetermined by the requirement of complete anti-correlation for given instrument settings, which is usually guaranteed and agreed upon through the complete experimental design of both wings. One must chose one instrument setting in one wing as gauge (like

the meter measure in Paris) and relate the gauge of the other wing and the so called Bell angles θ, by the convention that for maximum anti-correlation we have $\theta = 0$. Bell and his followers emphasize the free will of the experimenters to choose the instrument settings independently, randomly and swiftly in both wings. However, the spin-gauge cannot be chosen freely and independently as will be further discussed in detail. Oaknin [5] has given careful considerations to these facts and has offered a relativistic resolution for the EPR-paradox. (See also his detailed further explanations in his recent paper [6].)

Bell's original paper does not discuss photons but instead spin $\frac{1}{2}$ quantum entities (such as entangled electrons) that propagate toward Stern-Gerlach magnets instead of Wollaston prisms (cubes). The x, y - and x', y' -coordinates are again perpendicular to the propagation direction. Unit vectors that characterize the gradient of the magnetic field are introduced and typically denoted by **a**, which points in the positive *x*-direction in wing 1 and **b**, which points in the *x'*-direction in wing 2.

3. Bell's Functions and Their Relation to Einstein's Separation Principle

Bell attempted to describe EPRB experiments by introducing in each wing functions with a domain of variables that respect Einstein's separation principle. The co-domain or range of Bell's functions describes the outcomes for the measurements of entangled pairs. The mathematical physics of these functions is required to agree with the results of quantum mechanics and/or those of actual experiments.

In order to understand Bell's task, we need to explain the precise meaning of Einstein's separation for EPRB experiments and for Bell's functions that describe them.

Consider the two wing experiments described above and the case of maximal anti-correlation of the outcomes for entangled singlet pairs and assume that the measurement equipment of the two wings is spatially separated so that light would take considerable time (say a millisecond) to propagate from the place of measurement in wing 1 to the place of measurement in wing 2. Then, according to Einstein's relativity, whatever instrument setting is used in one wing and whatever is measured within about a millisecond cannot affect the outcomes of the measurement in the other wing by any transmission of information between the two wings. The measurement stations are information-separated in this way and this fact is called Einstein's separation principle.

Why did Einstein's separation principle enter the discussions of EPRB experiments? Because of the following strange fact. At the instrument settings of maximum anti-correlation, the outcome in wing 1 (e.g. "horizontal" or "vertical") is always anti-correlated to the outcome in wing 2 (which then is "vertical" or "horizontal", respectively *i.e.* the opposite) How can that be? If we toss a coin in wing one and obtain heads, the coin-toss in wing 2 must give tails. Einstein

claimed that this would only be possible if instantaneous influences are exerted from the measurement in one wing to the measurement of the other or, alternatively, there need to exist elements of physical reality (possibly "hidden" to us) that determine the measurement outcomes. Actually, one can easily agree with Einstein, when magnets are involved in the measurements. If coins with little magnets of opposite polarization are sent out from the source as entangled pairs to the measurement stations then it is fairly easy to imagine that one falls on head and the other on tail and that can also happen randomly. The hidden magnets in the coins are then Einstein's elements of physical reality.

The problem is, of course, that photons do not have little magnets inside and the elements of physical reality may be very complex and describe the dynamics of the photon equipment interaction etc. Furthermore, the experimental results need to be explained for all instrument settings and not only for maximal anti-correlation. Therefore, Bell attempted to develop a theoretical model for EPRB experiments in a more general fashion.

Bell introduced functions $A(...)$ for wing 1 and $B(...)$ for wing 2 with the following properties of the variables in the domain of the functions: one of the variables in each function is the magnet "setting", usually denoted by **a**, **b** or **d** in wing 1 and by **b** or **c** in wing 2. All other magnet settings (or equivalent settings of the Wollaston prisms) are possible but not considered in the following. The functions A and B contain each another special "variable" λ that represents the entangled pair. Bell states:

"… λ stands for any number of variables and dependences thereon of A and B are unrestricted."

In other words, λ may represent actually a whole set of variables that describe Einstein's elements of physical reality involved in the measurement of the entangled pair. Considering Einstein's theoretical thinking, such a set of variables might contain elements of the space-time continuum such as the measurement time t_m, because of possible correlations of the measurement dynamics that occurs in the two wings. Of course, such dynamic correlations may be described, in special cases, by a phase relationship between the two wings; the phase being again an element of a continuum.

The set of variables represented by λ is thus very general. In essence the set λ may contain any number of variables, but it must never contain a variable representing the instrument settings of the other wing, because these cannot influence the measurement outcomes in a given wing. In fact, to exclude instantaneous influences at a distance from the theory it is necessary and sufficient that the set λ in the domain of the function A in wing 1 must not contain any instrument-setting-variable from wing 2 and vice versa. With the given definitions, one can, mathematically speaking, use variables λ that are independent of the instrument settings of both wings. The elements of reality that λ represents may actually include local equipment interactions and, therefore, acquire some dependency on the local instrument settings. Mathematically, however, these

local settings are included already in the domain of Bell's functions anyway.

(Note also that there is a slight inconsistency in Bell's notation inasmuch he uses fixed settings **a**, **b** etc. in the domain of his functions while λ is standing for a variable and only in some instances for the value of this variable (a given element of physical reality). We leave Bell's notation as it is and just make sure that this inaccuracy does not introduce a bigger mistake.)

To illustrate by an example, allowed functions in wing 1 are $A(\mathbf{a},\mu,t_m,\cdots)$ and $B(\mathbf{b},\mu,t'_m,\cdots)$ in wing 2, respectively, where μ represents some element of physical reality related to information about the photons from the source. The measurement times t_m and t'_m may appear in the function domain because of possible time-like correlations of the dynamic photon-Wollaston or electron-Stern-Gerlach interactions in the respective wings. A and B are then, in Kolmogorov's framework, the functions (random variables) of bi-variate stochastic processes.

The values of the functions *i.e.* their co-domain are the detector registrations for the given measurement times, e.g. $D_v^1(t_m)$ or (exclusive) $D_h^1(t_m)$ in wing 1 and $D_v^2(t'_m)$ or (exclusive) $D_h^2(t'_m)$ in wing 2.

4. Logical Problems with the Applicability of Bell-Type Inequalities to EPRB Experiments

4.1. Problems with Bell's Function-Domain

One of the inaccuracies related to Bell's work (that Bishop Berkeley would have criticized) is the simplistic assumption that λ represents just a finite number of elements of physical reality like coins that are sent out by the source and registered by a detector-click after passing some "evaluation" equipment. If that were the case, we could surmise that for very many measurements with given setting pairs $(\mathbf{a},\mathbf{b})$, $(\mathbf{a},\mathbf{c})$ or $(\mathbf{b},\mathbf{c})$ etc. each setting pair must encounter about the same elements of physical reality that are essentially randomly sent out from a source. In other words, the expectation value of outcomes for given instrument setting pairs would just be an average of the function outcomes for these same elements λ. This simplistic assumption together with Bell's choice of co-domain or range of the functions (see next) leads immediately to Bell's inequalities. But do these inequalities then have anything to do with actual EPRB measurements?

4.2. Problems with Bell's Function-Range

A second suspicious assumption of Bell is an oversimplification of the co-domain or range of his functions. He innocuously introduces the value of +1 for a "horizontal" result and −1 for "vertical", respectively. (Or equivalently for spin $\frac{1}{2}$ entities Bell uses +1 for an "up" deflection by the Stern-Gerlach magnets and −1 for "down".) This use of the same two integer numbers for all possible instrument settings has far reaching consequences (see beginning of next

section), because it implies (without any justification) that the function-values follow the mathematical rules for the integers +1 and −1. Christian has emphasized repeatedly (most recently in [7]) that Bell oversimplified the range of his functions.

5. Bell Type Inequalities and Algebra

With the above assumptions we may deduce from the algebra of the integer numbers +1 and −1 that:

$$A(\mathbf{a},\lambda)\big(B(\mathbf{b},\lambda)-B(\mathbf{c},\lambda)\big)=A(\mathbf{a},\lambda)B(\mathbf{b},\lambda)\big(1-B(\mathbf{b},\lambda)B(\mathbf{c},\lambda)\big) \tag{1}$$

Note that the λs of the pair with a product of two B functions correspond necessarily to different entangled pairs (if we wish to compare Equation (1) with actual experiments). Furthermore, if we define the instrument settings for maximal anti-correlation as equal, use accordingly $B=-A$ and take the absolute value, we obtain:

$$\left|A(\mathbf{a},\lambda)B(\mathbf{b},\lambda)-A(\mathbf{a},\lambda)B(\mathbf{c},\lambda)\right|=1+A(\mathbf{b},\lambda)B(\mathbf{c},\lambda), \tag{2}$$

an equation that now features the $A\cdot B$ pair instead of the $B\cdot B$, suggesting that the outcome-pair now does correspond to the measurement of an entangled pair. The triviality of Equation (2) has persuaded many to believe that Bell's inequality (which almost immediately follows from it) is a simple consequence of algebra. A moment of reflection, however, shows that the algebraic operations of Equations (1) and (2) that Bell used after his Equation (14) are physically speaking not trivial at all and require extensive justification.

Bell further assumes the existence of a single common probability density, which might be appropriate if λ just represented a finite number of elements of physical reality; such as 20 fair or not so fair coins. However, Bell did claim the generality of λ and that λ may represent a whole set of variables, including measurement times. How, then, can all of these physical variables and the instrument setting variables have the same probability density?

Only with precisely one given probability density for all function products ($A\cdot A=-A\cdot B$), do we obtain Bell's inequality: Averaging over this single probability density $\rho(\lambda)$ and noting that the absolute value of a sum is always smaller than or equal to the sum of the absolute values, one obtains the expectation value E for the function products with Bell's instrument-setting pairs $\mathbf{a},\mathbf{b}$, $\mathbf{a},\mathbf{c}$ and $\mathbf{b},\mathbf{c}$:

$$\left|E(\mathbf{a},\mathbf{b})-E(\mathbf{a},\mathbf{c})\right|\le 1+E(\mathbf{b},\mathbf{c}), \tag{3}$$

where E may, in general, be expressed by a Lebesgue integral over the product $A\cdot B$ of Bell's functions.

6. Bell Inequalities and Probability Spaces

It was soon realized [8] [9] [10] that Bell's inequality could be derived with only one necessary and sufficient powerful condition: All functions appearing in

Bell's inequality are random variables on one given probability space in the sense of Kolmogorov. This fact makes it also easy to extend the validity of the inequality to a countable infinite number of the λs. However, two major problems still exist with the application of Bell's inequality to actual experiments and the results of quantum mechanics.

First, it turns out (shown by a theorem in [9]) that the measurement time and the instrument settings cannot be random variables on one common probability space. The reason is simply that we cannot have two different instrument settings in a given wing at the same measurement time. The proof of this fact is simple and may have been known to Einstein when he enunciated: "Gott wuerfelt nicht".

I would also like to point to the recent work of Khrennikov connecting quantum probabilities and classical conditional probabilities [11]. Novel arguments about the necessity of using different probability spaces for different setting pairs have been put forward (within the framework of quantum mechanics) by Cetto, Valdes-Hernandez and Pena [12].

Second, there exists a problem with the spin-gauges. How can one logically deal with two different "horizontal" and "vertical" directions in each given wing, ($(\mathbf{a},\mathbf{b})$ in wing 1 and $(\mathbf{b},\mathbf{c})$ in wing 2), and, in addition, regard all the measurement outcomes in the above equations equal to +1 or to −1 independent of how "horizontal" or "vertical" are globally defined for a given pair measurement?

7. Wigner Inequalities and Set Theory; Selecting Global Subsets

Eugene Wigner [3] improved Bell's treatment [1] significantly by using set theory and considering certain subsets of the measurement outcomes. Instead of using outcomes +1 and −1 as Bell does, Wigner relies only on the judgement of equal **e** and not-equal **ne** for the pair measurement outcomes of a given instrument setting pair.

The actual value and physical nature of the co-domain (integer numbers or "up/down") is thus of no concern, for we need to have only a judgement of "equal" or "not-equal". Wigner also noticed that, for the purpose to derive a Bell-type inequality, it is sufficient to determine the number of (possible) outcomes that are either **e** or **ne** for a given pair of instrument settings. He, therefore, just counted the number of equal and not-equal outcomes separately for each of the 3 pairs of Bell's instrument settings $(\mathbf{a},\mathbf{b})$; $(\mathbf{a},\mathbf{c})$ and $(\mathbf{b},\mathbf{c})$ and derived his Bell-type inequality using these numbers. The spin-gauge may also be (and must be, in principle), chosen as separate and different for each such given pair of instrument settings. Note, as a preview of the more detailed discussions below, that the judgement of **e** or **ne** is a global one that determines the relative outcomes and the frequency of them in both wings. Their frequency may, in general, depend on the instrument settings of both wings, which may lead to a

violation of the Bell-Wigner inequalities.

Bell [1], Wigner [3] and later d'Espagnat [13] do use, however, a single common probability space for all Bell instrument setting-pairs and the **e**-, **ne**-subsets, which enforces their inequalities. They do offer a weighty argument for this choice: If we consider a triplet-set of measurements (instead of the 3 pair measurements), we may write all possible measurement-outcomes in terms of triples:

$$A(\mathbf{a},...)A(\mathbf{b},...)A(\mathbf{c},...), \tag{4}$$

and using $A=-B$ automatically obtain one common joint-triple probability space that lets us construct the Bell-Wigner inequalities (think of the frequency interpretation of probability).

This last step, however, presents again major problems. First, a triplet is not measured in EPRB experiments, but we deal with entangled pairs. We, therefore, run again into the problem of the assumption of all equal λs, which cannot be justified if λ encompasses variables related to a continuum. We deal then with more complex bi-variate stochastic processes. One for each of the instrument setting pairs.

This problem is easily understood from the following example. Consider the three instrument setting-pairs of Bell and let λ be the measurement time. I assume, for simplicity, the measurement times in the two wings to be equal for a given instrument setting pair and given entangled pair and thus have the following possible measurement outcomes (Bell function-pairs):

$$A(\mathbf{a},t_m)A(\mathbf{b},t_m);A(\mathbf{a},t_n)A(\mathbf{c},t_n);A(\mathbf{b},t_k)A(\mathbf{c},t_k) \tag{5}$$

here we let $m=1,2,\cdots,N$ further $n=N+1,N+2,\cdots,2N$ and $k=2N+1,2N+2,\cdots,3N$, with N indicating a large number of measurements. Then, because all the outcomes for each of the function pairs may be either equal **e** or not-equal **ne**, we have $2^{(3N)}$ possible different combinations of **e** and **ne** outcomes. Bell's inequality, however, is based on the assumption of equal λs instead of different measurement times and considers, therefore, only $2^{(2N)}$ possible combinations of **e** and **ne** outcomes. This leaves us with $2^{(3N)}-2^{(2N)}$ of these possible **e** and **ne** outcome-combinations that may contribute to violations of the Bell-Wigner inequalities. (Actually, things are not quite as drastic as suggested by these considerations, because Wigner needs for his inequality only the fraction of the number of **e** and **ne** outcomes. Considering only combinations that change that fraction, one obtains $(N+1)^3$ different combinations if we include measurement times taken out of a continuum, as compared to N^2 for a finite (countable) number of the λs [14].)

Thus, the number of combinations of equal and not-equal outcomes that may violate the Bell inequality (given general time dependent functions as they are usual in physics) is vastly larger than the number of combinations that necessarily obey Bell's inequality; which is negligibly small for large N. If one would bet on the odds that any experimental sequence of EPRB-type experiments (de-

scribed by time dependent functions) violates the Bell inequality, one could bet with great certainty that it will. It is thus not surprising that so many different quantum experiments violate Bell-type inequalities.

As indicated above, there is also the other, even more disturbing, restriction and inconsistency in the treatment of Wigner that was later taken over by d'Espagnat [13]: they use only one common probability space, an assumption that is incorrect if one does not deal with Bell's λs, but instead with the subset of those λs that lead either to **e** or **ne** outcomes *i.e.* with $\lambda_{\mathbf{e}}$ and $\lambda_{\mathbf{ne}}$, respectively. Exactly how many of these "global" entities will lead to **e** or **ne** outcomes in the respective wings does in general depend, as I will demonstrate below, on the instrument settings of both sides without any involvement of instantaneous influences at a distance.

8. Alice and Bob to the Rescue: The Bell Game

The many objections related to infirmities of the derivation of Bell's inequality (and a large number of similar types of inequalities and equalities) made it desirable to find some way of stating Bell's findings crisply and without any possibility of objection. This was accomplished by putting the Bell inequalities themselves into the far background and just using their power of contradicting the quantum mechanical result. In this way Bell conceived a theorem for "local" theories, which is mathematically and physically always true, but only for a certain definition of "local". It was formulated by Bell in his following statement [15]:

"But if (a theory with … variables λ) is local it will not agree with quantum mechanics, and if it agrees with quantum mechanics it will not be local. This is what the theorem says."

Bell's definition of "local", however, does not only mean that the Einstein separation principle is strictly valid but adds additional, physically not necessary (actually physically impossible), conditions as explained below for the "Alice-Bob-game" or "Bell game", which was divined by Bell's followers and science writers.

Indeed, the above statement of Bell, with the addition of the "Alice-Bob-local" definition, is unassailable. As we will see, the theorem so stated is true and needs no mathematics to prove it. The problem with the theorem so stated is of epistemological nature: there exists no Alice-Bob-local physical theory of spatially distant and correlated measurements; not Einstein's relativity, not classical mechanics, not quantum mechanics, not any non-trivial theory as we will see momentarily.

What is the Alice-Bob-locality about? Alice and Bob are the experimenters in the two EPR wings and know only their own instrument setting and not that in the other wing during the measurements, because of the random switching of the instrument settings before a given measurement. It is postulated by the followers of Bell that Alice and Bob must be able to find a "local theory" that gives

them the outcome for Bell's functions as soon as they obtain an element of physical reality λ that is sent to them from a source. This "local-theory-game" is, as we will see, impossible to play and thus the Bell theorem is proven. Extreme non-localities, on the other hand, make it easy to find violations of Bell's theorem: just include a setting variable from the other side in Bell's functions (and use e.g. $A(\mathbf{a},\mathbf{b},\lambda)$) and Bell's inequalities are easily violated (spooky influences yield any desired results).

To understand why the "Alice-Bob-game" (also called by many "the Bell game") cannot be played and why the corresponding definition of "local" is too narrow and physically inappropriate, we have to realize that the experimenters must use non-local or global knowledge throughout their experimental design, during the measurements and after the measurements are finished and evaluated.

Much of the knowledge that Alice and Bob need about the other wing, can be gained by the design of the global experimental set-up before any of the actual measurements are done. Alice and Bob need to know that they measure the appropriate counterpart of an entangled pair. This is commonly accomplished by careful determination of the measurement times in both wings and requires the additional assumption that the measurement time does not depend on the equipment setting (which we may accept here without consequences for the following reasoning). The global gauge for all given instrument setting pairs and the settings for maximal anti-correlations as well as the connection of the coordinate systems of Alice and Bob, may also all be fixed before the actual measurements begin.

During the measurements, Alice and Bob do not have information about the global gauge that is relevant for a particular pair-measurement, because that gauge depends on the instrument setting of both wings and these instrument settings are randomly switched. It is, therefore, impossible for them to find a local theory that describes measurement outcomes violating Bell-type inequalities at this stage of the experiment. I have pointed out in a previous publication [16] that such requirements for a local theory would not permit any theory of relativity, which necessarily works with systems as seen relative to other systems, a fact that has been discussed by Oaknin [5] in great detail. Alice and Bob certainly cannot choose, at this stage, the Wigner subsets that are the key for finding violations of Bell-type inequalities and require the knowledge of the global gauge and instrument settings. The illustration in section 9 gives further reasons why Alice and Bob cannot play the Bell game.

The information about Wigner subsets is found by Alice and Bob only after all measurements are done: they assemble the Wigner subsets from the global data to prove the experimental violations of Bell-type inequalities by counting the equal **e** vs not-equal **ne** outcomes for the Bell-pairs of instrument settings. In this way they assess outcomes as seen relative to the other wing. The fact that this relativity is introduced after the measurement run does not make a differ-

ence to the fact that the statistics of the outcomes is determined by non-local means.

I have sometimes been asked by Bell's followers: "Nature can play the Bell game and finds measurement outcomes during the experimental runs, why can you not do it?" We see from the above that nature cannot play that local game, because it involves a global experimental design, a global gauge and a global assembly of Wigner subsets with the assessment of the frequency of relative outcomes in the two wings.

The theoretician who develops a theory about the Wigner subsets for **e** and **ne** outcome pairs, needs to have, as a minimum, that same global knowledge that the experimenters have and use to produce and finally collect their data as already outlined in [16]. These latter global facts and knowledge may not only reasonably be used by the theoretician, but must be used when probabilistic theories are invoked, because the probability spaces of subsets may depend on these global facts: when we ask the question of how many **e** and **ne** pairs we have for a given instrument setting pair, we do address a global fact for which only a globally valid theory and gauge can account. In fact we ask what is the outcome relative to the outcome in the other wing and we lose the possibility to treat the two wings independent of their respective measurement outcomes. Wigner improved Bell with respect of the generality of the function range (co-domain) but had to deal, in return, with the numbers of λ that correspond to **e** and **ne** outcomes and these numbers (or their frequencies of occurrence) may depend on the instrument settings (gauge) of both wings because they depend on the relative measurement outcome of the other wing. This latter fact will be shown in more detail in section 9.

In this connection it is also important to remember that non-locality by instantaneous influences at a distance requires a specific inclusion of the instrument-settings of the other wing in the domain of Bell's functions. Such inclusion is neither present nor required at all in the following illustration.

9. Wigner-Subsets, Spin-Gauge and Probability-Spaces: The Role of Free Will and Randomly Switched Instrument Settings

The gauge for the spin ("horizontal/vertical" in a certain coordinate system) can be chosen freely in only one wing of the EPRB experiment. The gauge in the other wing is determined by the requirement of complete anti-correlation, which also determines the x, y coordinates on both sides. The outputs of the Wollaston prisms (Stern-Gerlach magnets) are then anti-correlated by definition for the same x, y coordinates, as they should be for the singlet entangled pairs that we consider.

This latter important point and its full consequences were not realized by Bell and his followers. They claim that the settings on both sides are chosen by the free will of Alice and Bob. Naturally, even the free will cannot turn instruments faster than the speed of light c, which is the reason why measurement times and

instrument settings cannot be defined on one common probability space [9]. Furthermore, the gauge of the spin must be globally well-defined also in the second wing, once chosen in the first. Detailed considerations about using consistent spin-gauges were presented by David Oaknin [5]. They were also discussed independently and from a different vantage point by this author [17] and in an early stage with collaborators [18]. It is important to realize that there are certain restrictions for the choice of gauge, which have been discussed in all generality by Oaknin [5]. For our purposes here it is sufficient to adopt a definition of gauge corresponding to one given instrument setting (say in wing 1) and to connect this gauge to wing 2 as described above and below.

We may rotate the Wollaston prism in wing 2 freely, for example by an angle θ around the z-axis and replace thus the measurement directions x, y by the rotated x', y' coordinates in wing 2. Note, however, that we must now label the measurement data in wing 2 by the angle θ between x and x', if we wish to say anything about the Bell-correlations of the two wings including complete anti-correlation for the x, x' settings. Oaknin [5] states: "… only their relative orientation (referring to x, x') is a physical degree of freedom." If we do not relate the wing 2 instrument settings and wing 2 gauge to that of wing 1, we naturally cannot speak about correlations. Remember also that in actual EPRB experiments the determination of the Bell angles is usually done by the experimenters before the actual measurement-runs, while the choice of measurement sequences and setting pairs is made after the experimental runs are finished; when the Wigner subsets are collected mostly based on measurement times.

The rotation of the Wollaston in wing 2 by θ results, of course, in measurement outcomes that are different from complete anti-correlation. The corresponding correlations of the outcomes in this new situation are still determined by the physical law that governs the interactions of the photons and the Wollaston prisms and the possible measurement outcomes may, therefore, be different in wing 2 in a variety of ways.

The following illustration by an example from classical mechanics is designed to explicitly demonstrate the associated possible changes of probability spaces by rotations of the Wollaston prism and by subsequently choosing Wigner-type subsets of **e** and **ne** outcomes. We will see that such rotation changes the Kolmogorov probability spaces for the **e** and **ne** outcomes and involves θ or functions of θ. It will also become obvious that these changes of the probability spaces with θ have nothing to do with instantaneous influences at a distance in spite of the fact that θ depends on the instrument settings of both sides.

9.1. Illustration of Global Correlations between Classical EPRB-Type Measurement Pairs without Violation of Einstein's Separation Principle

This oversimplified illustration (per se) is not invalidating Bell's inequality, because it is linear in all its variables. It is demonstrating in the most elementary way, however, that a dependence of certain probabilities on θ has nothing to

do with influences at a distance but arises naturally from local factors and the global experimental design and gauge.

Consider two macroscopic rods parallel and next to each other, one with a red top and blue bottom section, the other with a blue top and red bottom section. The top of one rod is always next to the bottom of the other. The center of these rods is on the z-axis of a coordinate system and the rods are always oriented perpendicular to the z-axis. Assume further that these rods are emitted from a source. The direction of the emission is random: either the first rod propagates into wing 1 and the second into wing 2 along the z-axis or vice versa, both rods exhibit the same angle ϕ, which is measured from the *x*-axis of the x, y-plane and ranges in value randomly between 0 and π.

In both wings we have instead of the Wollaston prisms simple color detectors that indicate the result "horizontal" if and only if for any ϕ between 0 and π the color is red, while they indicate "vertical" if for any such ϕ the color is blue. We can see that this simple experiment will always result in complete anti-correlation without any suspicion of instantaneous influences at a distance. There is, of course, global knowledge and fact involved in the gauge, the physical conservation law that guides the rods toward the detectors and how the rods interact with the detectors locally but in a correlated way, because of the overall design of the experiment.

Assume now that many such rod-pairs are emitted from a source. We use the same local rules for evaluating the red and blue rod-sections, but we rotate the coordinate system and color detectors in wing 2 by an angle θ that turns x into x'. We denote the angles of the rods with the x'-axis by ϕ' and determine the new "horizontal" and "vertical" outcomes in wing 2 according to the colors red or blue, respectively. Now, however, we do this for ϕ' in the x', y'-plane with values between 0 and π.

After performing many measurements, we select in wing 1 the subset of rods with the exclusive outcome "horizontal". We then ask the question, which fraction of the corresponding rods in wing 2 is "vertical" and which fraction "horizontal". The "entangled" pair of rods is determined by the angle ϕ and the fact that $\phi' = \phi - \theta$, another needed global knowledge. Because ϕ is random between 0 and π, one obtains $1 - \frac{\theta}{\pi}$ for the "vertical" fraction of rods and $\frac{\theta}{\pi}$ for the "horizontal" fraction, respectively. Thus, the probability (using the frequency interpretation) of the "horizontal" and "vertical" outcomes depends on the angle θ between the two coordinate systems.

Alice and Bob, however, could not possibly guess that simple result; they cannot know about the instrument setting in the other wing during the measurements and thus also do not know the angle θ. The probability measure $P(A = B)$ for the outcomes to be **e**, equals the probability that $0 \le \phi \le \frac{\theta}{\pi}$, which equals $\frac{\theta}{\pi}$. As is evident from the simplicity of the model, the dependence of the probability measure on the angle between the instrument settings in the

two wings does not indicate any implausible non-locality and certainly not any influences at a distance but arises from the global factors of the experimental arrangement, gauge and choice of Wigner subsets.

This means that Wigner-subset probability-measures for Bell's function products and the numbers of **e** and **ne** outcomes may depend on the instrument settings of both sides, without involvement of spooky non-localities.

As a major corollary one can state that a global statistical result, obtained from many measurements at separate locations for correlated information packages and correlated measurement times, may depend on non-local variables such as θ. The global statistical result may reflect measurement arrangements of all of the separate locations, even if those arrangements do not influence each other and are unknown to anyone controlling the local measurement events. It is this corollary, which permits us to use the space-time system to exorcise spooky influences in complex situations if we choose to do so.

Note that these statistical properties of subset probability-measures do not imply that Bell's function-domain contains a variable corresponding to the equipment settings of the other wing. Nor does the random angle ϕ, which corresponds to Bell's λ (except that it is chosen out of a continuum of angles) depend on any instrument settings. It is only the probability for **e** and **ne** outcomes that depends on the instruments of both wings. Nor is the free will of Alice and Bob to choose any angle θ restricted in any way (super-determinism). They must choose, however, the spin-gauge of both wings consistent with the settings of complete anti-correlation and cannot decide that fact by their separate free will.

9.2. The Importance of Subset Selection and the Problem with Random Switching

Fast random switching of the instrument settings (and thus of θ) are declared by Bell's followers to be the vade mecum for proving Bell's theorem, because it makes it impossible to play the Alice-Bob-local game (the Bell game). Neither Alice nor Bob know θ and the gauge that is chosen in the other wing, because that gauge depends on the (rapidly switched) instrument setting. Their choice of the outcome-value for Bell's functions (of $A(\mathbf{a},\lambda)$ by Alice and $B(\mathbf{b},\lambda)$ by Bob, after they receive a value of λ) is, therefore, meaningless at the time the measurements are performed. They both do not know the global gauge at this point. This knowledge is only acquired by them when they select the Wigner subsets after the measurements are completed.

Random switching of the measurement settings on both sides does involve random changes of the spin-gauges and of the probability measures for the **e** and **ne** outcomes of selected subsets. It is, therefore, nonsensical to require that Bell's functions and the ordering of their outcome-values into subsets a la Wigner are related to only one common probability space. The probabilities $P(A=B)$ and $P(A\neq B)$ depend in our (classical mechanics) illustration on θ, which is ran-

domly changed by randomly switching the instrument settings. The angle ϕ (chosen out of a continuum), which corresponds to Bell's λ emanating from the source, does not depend on the instrument settings as dictated by Einstein's separation principle. However, the frequencies of **e** and **ne** outcomes may depend on the instrument settings (and do depend on them in our illustration), because of the underlying physical law, global experimental arrangement and global gauge.

These facts make us appreciate Einstein's view (in his discussion with Heisenberg) that the theory codetermines what can be measured. The experimenters need to know about Wigner subsets in order to find out whether or not Wigner's inequality is obeyed by their data. The theoretician needs to help with the construction of the global design and needs to provide a consistent global gauge. The requirement that Alice and Bob should be able to find a theory at a certain point of the experimental procedure at which they have no idea of the global gauge and other factors, appears in this light as a crucial mistake.

Quantum mechanics also gives results for precisely one given instrument setting in each wing. That setting pair determines the operators that act on the designated states of a product Hilbert-space. The instrument setting thus determines the spin operators and defines the gauge through the eigenvalues and eigenvectors of these operators. The preparations of particles and their quantum states (on which the spin-operators act), determine the precise division into different subsets of the possible outcomes corresponding to **e** and **ne** values. These different subsets are, however, not necessarily defined on one common probability space but involve, in general, different probability spaces for different instrument setting pairs.

10. Explicit Statistical Interpretation and Model for EPRB

How can one model the general probability measures for the actual singlet-entangled-pair measurement outcomes, without invoking any inappropriate nonlocal influences? How can one exclusively use the data and invoke a physical law that explains the probability measures without instantaneous influences at a distance?

I specify our considerations to the case of entangled photon pairs and measurements with Wollaston prisms and also consider only Wigner's set theoretical approach, which means we need to determine only the number of **e** and **ne** outcomes for a given instrument setting pair say **a** in wing 1 and **a** as well as **b** in wing 2. This use of Wigner-type subsets, which are selected by the theoretician after all measurements are completed, is crucial.

We follow the above illustration and collect from all data, as a first step, the M "horizontal" outcomes of wing 1, as well as the corresponding measurement outcomes for the entangled photons in wing 2. For $\theta = 0$ (setting pair $\mathbf{a},\mathbf{a}$) all the outcomes in wing 2 must be "vertical" because of maximal anti-correlation. For $\theta \neq 0$ and Wollaston prism in wing 2 set to **b** (the x axis rotated by an an-

gle θ to form the *x'*-axis), one naturally expects a Malus type law at work (see e.g. The FeynmanLectures on Physics III), which results in about $M(\sin\theta)^2$ "horizontal" outcomes, because that is what is natural for a system of entangled pairs that had shown all "vertical" outcomes before the Wollaston prism was rotated by θ. The "vertical" outcomes in the rotated system are then $M - M(\sin\theta)^2 = M(\cos\theta)^2$. These results are also expected by quantum considerations as shown in the Feynman lectures.

As in our illustration above, the use of the angle θ involves no "illegal" nonlocality. We have used a Malus-type law for the probabilities of "horizontal" and "vertical" outcomes.

Symmetrically, for the set of all "vertical" outcomes in wing 1, we obtain in wing 2 about $M(\cos\theta)^2$ "horizontal" outcomes and $M(\sin\theta)^2$ "vertical" outcomes. One easily obtains then the number of equal outcomes to be $\mathbf{e} = 2M(\sin\theta)^2$, while the number of non-equal outcomes is $\mathbf{ne} = 2M(\cos\theta)^2$. Thus, we have for the difference in the outcome probabilities (frequencies):

$$E(\mathbf{a},\mathbf{b}) = -\left((\cos\theta)^2 - (\sin\theta)^2\right) = -\cos(2\theta) \tag{6}$$

Nothing in the procedure depends on the distance of the Wollaston prisms. Nor do we need to involve any instantaneous influences at a distance.

We got around the instantaneous changing of a "state" by avoiding, a la Einstein, relations of state concepts to single objects and instead using all horizontal outcomes in wing 1 as a subset and obtaining the correlated subsets in wing 2 by the probabilities as obtained from a Malus type law.

In contrast, the quantum-interpretation of Bell's followers maintains that (because the outcomes that we consider in wing 1 are all "horizontal") we are dealing in wing 1 with measurement of a single "horizontal" quantum state and only after these measurements do we know that the state in wing 2 must be a vertical quantum state and instantaneously so.

We have, with the outcome-subset treatment, not achieved any progress or improvement of the results, but have avoided instantaneous changes of single-object quantum states. In other words, we can clearly avoid any hint of instantaneous influences in this model for EPRB experiments. The major novelty is that certain subsets of data-pairs may necessarily be defined on different probability spaces. Quantum mechanics avoids dealing with this complication by only working with probability amplitudes that lead to the final subset probabilities by Born's interpretation. Andrei Khrennikov has analyzed the matter of statistical vs individual interpretations from a mathematical point of view and has contributed over many years to a large number of questions discussed above [19]

Would Einstein be satisfied with the model here presented? Not quite, because he preferred ab initio physical theories that do not invoke probability laws. However, if we do not wish to introduce a Malus type law of nature, then we must dive into the nitty gritty of the dynamic interactions of entangled pairs and measurement equipment and describe them in space-time. This presents a very complicated theoretical problem that may lead to the well-known infinite regress

that arises if the equipment is treated as a many body system. This difficulty is the great barrier for ab initio Bell-type models and this barrier has contributed to many frustrations in this area. It is identical to the similar barrier of quantum theory that usually does not and cannot supply an ab initio theory of the measurement instruments.

Of course, there are numerous other quantum experiments exhibiting the same features that I have just described for EPRB: they are much easier to explain with instantaneous influences at a distance than without those influences. In the final analysis, Einstein found the use of such influences spooky and spook must be exorcised. However, spook is sometimes "entrenched" in highly valued physical concepts and strategies of explanation, simply because it solves all problems easily.

Typical examples are experiments that can be explained with interference at distant regions of space, particularly interference of probability amplitudes. However, we encounter here again a suspect logic that Bishop Berkeley would have objected to: how can the interference of "something" destruct any results at distant spatial locations, where that "something" is, ontologically speaking, not present in the first place? In other words, the concept of interference at a distance involves instantaneous influences at a distance to start with. Instead of embracing such concepts and using them to unseat other logical explanations, work must be done, as above for the model of EPRB experiments, that permits to remove instantaneous influences at a distance and exorcise spooky effects. Such work is not going to be easy, as one can see from the years of controversy with the Bell theorem. However, only such dedicated work can lead us away from spooky influences and toward more Einstein-like theories.

11. Conclusions

It has been examined whether EPRB experiments can be reasonably modeled without the use of instantaneous influences at a distance and a way was found that is numerically identical to the quantum mechanical results but interpretational different. The concept of quantum-state as a description of a single entity has been avoided by using the description of subsets of entities interacting with instruments and resulting in subsets of data. A main difference between quantum and Kolmogorov probability concepts has been pinpointed in the treatment of such subsets of the data for measurement outcomes. These subsets correspond in quantum mechanics to different quantum states as well as operators, while in Kolmogorov's framework they correspond generally to different probability spaces.

I have also shown that Bell's theorem does not apply to such subsets of data-pairs of distant EPRB measurement outcomes, because the definition of "local" by Bell and followers in the Alice-Bob-game is too narrow. With that definition, the Bell theorem is indeed valid, but does not apply to any nontrivial actual experiments.

I have further shown that the Bell inequalities themselves suffer from a number of logical inconsistencies and lack of generality in their dealing with probabilistic concepts, inconsistencies that also cannot even be repaired by the more general set theoretic approach of Wigner and d'Espagnat.

I believe that my findings suggest that instead of embracing instantaneous influences at a distance in physical theories and in quantum mechanics, ways must be searched for that avoid the use of such concepts by the appropriate use of different subset-probability-spaces. Quantum mechanics, of course, has accomplished that division into sub-sets through quantum states and operators. However, the interpretations that link quantum states to single ontological entities lead to temptations of suggesting instantaneous influences, which may be avoided by careful additional explanations as presented above for the special case of EPRB.

Acknowledgements

I would like to thank Louis Marchildon, Hans De Raedt, David Oaknin, Juergen Jakumeit and Andrei Khrennikov for important and valuable discussions.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Bell, J.S. (1964) *Physics*, **1**, 195-200. https://doi.org/10.1103/PhysicsPhysiqueFizika.1.195

[2] Einstein, A., Podolsky, B. and Rosen, N. (1935) *Physical Review*, **16**, 777-780. https://doi.org/10.1103/PhysRev.47.777

[3] Wigner, E.P. (1970) *American Journal of Physics*, **38**, 1005-1009. https://doi.org/10.1119/1.1976526

[4] Weihs, G., Jennewein, T., Simon, C., Weinfurther, H. and Zeilinger, A. (1998) *Physical Review Letters*, **81**, 5039-5043. https://doi.org/10.1103/PhysRevLett.81.5039

[5] Oaknin, D.H. (2015) Solving the EPR Paradox: An Explicit Statistical Model for the Singlet Quantum State. https://arxiv.org/abs/1411.5704

[6] Oaknin, D.H. (2020) *Frontiers in Physics*, **8**, Article 142.

[7] Christian, J. (2018) *Royal Society Open Science*, **5**, Article ID: 180526. https://doi.org/10.1098/rsos.180526

[8] Hess, K. (2015) Einstein Was Right! Pan Stanford Publishing, Singapore. https://doi.org/10.1201/b16809

[9] Hess, K. and Philipp, W. (2005) *Foundations of Physics*, **35**, 1749-1767. https://doi.org/10.1007/s10701-005-6520-y

[10] Accardi, L. (1981) *Physics Reports*, **77**, 169-192. https://doi.org/10.1016/0370-1573(81)90070-3

[11] Khrennikov, A. (2015) *Foundations of Physics*, **45**, 711-725. https://doi.org/10.1007/s10701-014-9851-8

[12] Cetto, A.M., Valdes-Hernandez, A. and de la Pena, L. (2020) *Foundations of Physics*, **50**, 27-39. https://doi.org/10.1007/s10701-019-00313-8

[13] d'Espagnat, B. (1979) The Quantum Theory and Reality. Scientific American, New York, 158-178. https://doi.org/10.1038/scientificamerican1179-158

[14] Hess, K., De Raedt, H. and Michielsen, K. (2017) *Journal of Modern Physics*, **8**, 57-67. https://doi.org/10.4236/jmp.2017.81005

[15] Bell, J.S. (1976) *Epistemological Letters*, **2**, 2-7.

[16] Hess, K. (2019) *Journal of Modern Physics*, **10**, 1209-1221. https://doi.org/10.4236/jmp.2019.1010080

[17] Hess, K. (2018) *Journal of Modern Physics*, **9**, 1573-159. https://doi.org/10.4236/jmp.2018.98099

[18] Hess, K., Philipp, W. and Aschwanden, M. (1982) *International Journal of Quantum Information*, **4**, 585-625. https://doi.org/10.1142/S0219749906002080

[19] Khrennikov, A. (2008) *Entropy*, **10**, 19-32. https://doi.org/10.3390/entropy-e10020019

Journal of Modern Physics, 2020, 11, 703-711
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Melia's $R_h = ct$ Model Is by No Means Flat

Rainer Burghardt

A-2061 Obritz 246, Obritz, Austria
Email: arg@aon.at

How to cite this paper: Burghardt, R. (2020) Melia's $R_h = ct$ Model Is by No Means Flat. *Journal of Modern Physics*, **11**, 703-711.
https://doi.org/10.4236/jmp.2020.115045

Received: April 7, 2020
Accepted: May 15, 2020
Published: May 18, 2020

Abstract

With the support of numerous arguments, it has been shown that Melia's claim that his cosmological $R_h = ct$ model is flat and infinite is erroneous. In contrast, the model is positively curved, closed and, therefore, finite. With respect to results of Melia's model, it is identical to our Subluminal Model.

Keywords

Globally and Locally Flat Cosmos, Melia's $R_h = ct$ Model, Subluminal Model, Curvature Parameter

1. Introduction

In numerous papers[1], Melia has proposed a cosmological model that is flat and infinite and thus contains an infinite amount of matter. Matter, space, time, and infinity were created with the Big Bang. Melia called his model the $R_h = ct$ model, where R_h represents the non-comoving radial coordinate at the cosmic horizon of the expanding model and t the cosmic time, *i.e.*, the time in the system that comoves with the expansion. Evidently, the name of Melia's model comprises two variables belonging to two different coordinate systems.

Melia has an extensive set of astrophysical data and has demonstrated in a series of articles that this data can be best adapted to the $R_h = ct$ model as compared to other Friedman-Robertson-Walker (FRW) models. Moreover, his model provides an exact solution to Einstein's field equations while most FRW models do not. Therefore, Melia's model is significantly different from the standard FRW model, where pressure is applied by hand. As Einstein's field equations cannot fully determine the FRW models, it is necessary to introduce numerous parameters, namely, the Ωs and the deceleration parameter. These quantities must be determined using astrophysical data. However, for Melia's model, only one parameter needs to be determined. This feature explains why Melia's model

[1]Most papers by Melia and colleagues are listed in [1].

is favored over other models.

We proposed our Subluminal Model [1], which is positively curved and closed. The main aim of this paper is to prevent the assumption that galaxies in the universe have superluminal velocities. Surprisingly, Melia's $R_h = ct$ model and our Subluminal Model yielded the same final results. One obtains the same Friedman equation, the same EOS, *i.e.*, $\mu_0 + 3p = 0$, and a uniform expansion of the universe.

In the second section, we oppose the two models and show that both models, although derived in different ways, are identical. We collect the results from our earlier papers and comment on them. We believe that Melia's universe is also positively curved and closed. Thus, Melia's observational results are valid for our model as well. Nevertheless, we assert that Melia's geometrical interpretation of his model is erroneous.

2. The Question of Curvature

In the following section, we will demonstrate step by step that Melia's $R_h = ct$ model is not flat at all. We have referred to our earlier papers [2] [3] where we treated the same aspect of the problem. Here, we have summarized the results.

i) We favor the view that infinite universes, be they open, flat, or negatively curved, are ruled out as a way of describing Nature. This is because infinities are hard to imagine and because we want to avoid conclusions from Hubble's law that lead to acausality and contradictions to the special theory of relativity. This is one of the reasons we have rejected the geometrical interpretation of Melia's infinite model.

ii) An infinite universe has to expand to avoid Olbers' paradox. An infinite number of stars emit an infinite amount of light. Although the intensity of light decreases with $1/r^2$, the night sky will be as bright as our sun. In the case of expansion of the universe, distant stars run away and influence the $1/r^2$ law, thereby avoiding Olbers' paradox.

iii) According to Hubble's law, $v = Hr$, where H, as the Hubble parameter, associates the recession velocity v of the galaxies with the distance r of an observer. This law emerges from astrophysical observations. Evidently, in an infinite universe, the distance r can be chosen to be arbitrarily large, and the recession velocity may reach or exceed the velocity of light. Thus, the formation of galactic islands could be possible. However, no information can be exchanged between such galaxies. The laws of special relativity are inevitably violated in an infinite universe.

iv) The boundary where the recession velocity becomes superluminal is known as the cosmic horizon. Melia introduced such a horizon rather artificially. He, building on a flat universe, created an event horizon by comparing it with the Schwarzschild theory. He referred to Weyl's cosmological principle and Birkhoff's theorem. An enclosed mass $M = M(r_h)$ of a certain volume in the universe determines the Hubble radius[2] $r_h = 2GM/c^2$, leading to the relation

[2]Melia's variables R, t correspond to our variables r, t', which we have used in our earlier papers.

$r_h = ct'\left(R_h = ct\right)$. The Hubble radius is defined as the distance light has traveled since the Big Bang; t' represents the age of the universe and r_h the location at which the rate of expansion reaches the speed of light.

It should be mentioned that to define a cosmic horizon, the mass content of the universe is not mandatory. This shows the original version of the dS cosmos, which is empty and has a horizon.

v) In contrast, our Subluminal Model has a natural horizon. This model is based on the dS model, which can be geometrically represented by a 4-dimensional pseudo-hypersphere with a constant radius $\Re$ embedded into a 5-dimensional flat space. Expunging the condition $\Re = const.$, one can obtain our Subluminal Model.

The pseudo-hypersphere of the dS cosmos is usually described with pseudo-spherical coordinates $r, \vartheta, \varphi, x^4 = it = \Re i\psi$. Here,

$$r = \Re \sin\eta \tag{2.1}$$

is the radial coordinate and η the polar angle of the pseudo-hypersphere. Choosing an arbitrary point on the pseudo-hypersphere as a pole, *i.e.*, the location of an observer, the associated equator surface ($\sin\eta = 1$) has

$$r_h = \Re, \tag{2.2}$$

the maximal extension of *r*. This is the natural geometrical definition of the cosmological horizon and is equally valid for the Subluminal Model. It is the basic relation that connects the two models under consideration. Thus, the recession velocity is also limited via Hubble's law. As we have already shown in our paper [4], the geometric horizon can be reached by drifting galaxies only after infinite time, as experienced by the chosen observer. As observers can be fixed at any arbitrary point on the pseudo-hypersphere, each observer has an individual horizon.

vi) To examine the relationship between the two models in greater depth, let us revisit the abovementioned definition of Melia's cosmic horizon. Melia determined the Hubble radius with

$$r_h = \frac{2GM\left(r_h\right)}{c^2}.$$

Here,

$$M\left(r_h\right) = \frac{4\pi}{3}\frac{r_h^3}{c^2}\mu_0$$

is the mass enclosed by the sphere with radius r_h and μ_0 as the assigned mass density. Thus, with the aid of (2.2), we get

$$r_h = \sqrt{\frac{3c^4}{8\pi G\mu_0}} = \sqrt{\frac{3}{\kappa\mu_0}} = \Re. \tag{2.3}$$

This immediately results in

$$\kappa\mu_0 = \frac{3}{\Re^2}, \tag{2.4}$$

an expression derived in our Subluminal Model with geometrical methods.

The mass density decreases as the universe increases with the radius $\Re$. This and similar relations can also be found in Einstein's universe, Friedman's universe, and in the models of the dS family. However, this relation is missing in Melia's papers. As Melia's model is assumed to be flat, a familiar relation (2.4) cannot be derived within the framework of his model.

vii) Both models the $R_h = ct$ model and the Subluminal Model describe the relation between the non-comoving radial coordinate r and the comoving coordinate r' with

$$r = \kappa(t')r', \tag{2.5}$$

where κ is the time-dependent scale factor. We still need to demonstrate that the $R_h = ct$ model is compatible with the features of the curvature of the pseudo-hypersphere. With

$$r = \Re \sin\eta, \quad r' = \Re_0 \sin\eta, \quad \Re = \kappa\Re_0, \quad \Re_0 = const., \tag{2.6}$$

we can write the Hubble parameter with both the scale factor and the pseudo-hypersphere's radius of curvature as

$$H = \frac{\dot{\Re}}{\Re} = \frac{\dot{\kappa}}{\kappa}, \tag{2.7}$$

where $\Re_0$ is the radius of the curvature of the pseudo-hypersphere if it is calculated with the aid of comoving, *i.e.*, expanding rods, and thus appears to be a constant quantity for the comoving observer.

In addition, by solving the field equations of the Subluminal Model, we can obtain the mass density, the pressure, and the EOS as follows:

$$\kappa\mu_0 = \frac{3}{\Re^2}, \quad \kappa p = -\frac{1}{\Re^2}, \quad \mu_o + 3p = 0. \tag{2.8}$$

Both pressure and mass density are functions of the time-dependent radius of the universe.

viii) In the Subluminal Model, the recession velocity is *geometrically* defined by

$$v = \sin\eta = \frac{r}{\Re}. \tag{2.9}$$

Respecting (2.2), the recession velocity at the horizon is

$$v_h = 1,$$

the velocity of light in the natural measuring system. In addition, solving Friedman's equation, we can arrive at the following simple relation:

$$\dot{\Re} = 1. \tag{2.10}$$

With (2.7) one has $H = 1/\Re$, and using Hubble's law, we can confirm (2.9). The dot in Equation (2.10) denotes the derivation with respect to cosmic time. Thus, we can now recover Melia's fundamental relation $R_h = ct$ using the physical measuring system.

ix) The essential difference concerning the interpretation of the models is the

question of curvature, *i.e.*, the interpretation of the curvature parameter *k*. The line elements of both models in comoving coordinates are the same, and we can see from them that $k=0$.

In our Subluminal Model with the static dS metric as the seed metric has $k=1$. The geometry of the dS cosmos is interpreted as the pseudo-hypersphere and thus as a positively curved, finite universe. It should be noted that a coordinate transformation to the comoving system cannot change the curvature of the space. Due to this, we insist that the curvature of both models is positive and the universe is finite.

It is well known to gravitational physicists that a transition to a freely falling coordinate system does not change the geometrical structure of a model. Lemaître found such a transformation for the Schwarzschild model. Observers in a freely falling elevator tend to hover, implying that they are not exposed to gravitational forces. When writing the line element of the static Schwarzschild model in canonical form, the curvature parameter is $k=1$; however, in a freely falling system is $k=0$. This shows that *k* is not a reliable criterion for the curvature of space. In contrast, $k=0$ denotes that an observer is in free fall. Thus, it would be more convenient to call *k* a form parameter for a metric. In our paper [3], we have discussed this problem in detail and extended Einstein's elevator principle to cosmology. We conclude $k=0$ in Melia's model does not necessarily indicate the model is flat but rather indicates that the universe is in free fall.

x) In an extensively quoted paper [5], Florides discussed the relations between comoving and non-comoving coordinate systems for several cosmological models. In our papers [6] [7], we complemented the coordinate transformations of Florides with Lorentz transformations. From all these papers, it can be seen how the parameter *k* changes under coordinate transformations. In a table, we have provided an overview of cosmological and gravitational models in [8] and shown that *k* assumes rather individual values depending on the choice of coordinates for the line element. Therefore, the statement of numerous authors at the beginning of their articles that $k=(1,0,-1)$ denotes positively curved, flat, or negatively curved spaces is definitely wrong. Florides states that the only physical acceptable member of the dS family is the de Sitter cosmos, *i.e.*, the universe with a metric that transforms $k=1$ into $k=0$. This is the very metric we have used as the seed metric for our Subluminal Model.

xi) To determine the structure of the universe, we cannot rely on the parameter *k*. Instead, Einstein's field equations need to be solved and the geometrical properties of the given quantities studied. Unfortunately, several cosmologists tend to manipulate Friedman's equation without considering the remaining components of Einstein's field equations. This way, they propose new models, trying to explain dark matter or dark energy and other possible effects in cosmology. It can be said that such solutions are not exact solutions to Einstein's field equations.

A complete treatment of the field equation can disclose the geometrical structure of the model and determine the curvatures of space. We will demonstrate

this in a somewhat pedagogical manner. Starting with a simple 2-sphere embedded into a 3-dimensional flat space, the line element on this sphere can be expressed as follows:

$$\mathrm{d}s^2 = r^2\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\mathrm{d}\varphi^2 . \tag{2.11}$$

Here, r represents the radii of the greater circles and $r\sin\vartheta$ the radii of the parallels. From the line element, we can read the tetrads and calculate the Ricci-rotation coefficients[3]

$$A_{mn}{}^{s} = B_{mn}{}^{s} + C_{mn}{}^{s}, \quad B_{mn}{}^{s} = b_m B_n b^s - b_m b_n B^s, \quad C_{mn}{}^{s} = c_m C_n c^s - c_m c_n C^s$$
$$b_m = \{0,1,0\}, \quad c_m = \{0,0,1\}, \quad m = 1,2,3$$

Herein, the *curvatures* are defined by

$$B_m = \left\{\frac{1}{r},0,0\right\}, \quad C_m = \left\{\frac{1}{r\sin\vartheta}\sin\vartheta,\frac{1}{r\sin\vartheta}\cos\vartheta,0\right\}. \tag{2.12}$$

We see that the Ricci-rotation coefficients contain the curvatures of the sphere $\frac{1}{r},\frac{1}{r\sin\vartheta}$. It is easy to extend this method to higher-dimensional spaces.

xi) The dS model is based on the metric

$$\mathrm{d}s^2 = \Re^2\mathrm{d}\eta^2 + \Re^2\sin^2\eta\mathrm{d}\vartheta^2 + \Re^2\sin^2\eta\sin^2\vartheta\mathrm{d}\varphi^2 + \Re^2\cos^2\eta\mathrm{d}i\psi^2 . \tag{2.13}$$

It is the metric of a 4-dimensional pseudo-hypersphere

$$x^{a'}x^{a'} = \Re^2, \quad a' = 0',1',\cdots,4'$$

with a constant radius $\Re$ embedded into a 5-dimensional flat space, parametrized by

$$\begin{aligned} x^{3'} &= \Re\sin\eta\sin\vartheta\sin\varphi \\ x^{2'} &= \Re\sin\eta\sin\vartheta\cos\varphi \\ x^{1'} &= \Re\sin\eta\cos\vartheta \\ x^{4'} &= \Re\cos\eta\sin i\psi \\ x^{0'} &= \Re\cos\eta\cos i\psi \end{aligned} \quad .$$

Here, a' denotes the Cartesian coordinate system of the embedding space, where $x^{4'} = \Re\cos\eta\ i\ sh\psi$ is related to an imaginary dimension of space, the "Cartesian time." To understand the curvature problem, we can restrict ourselves to the greater circles of the spherical piece of the pseudo-hypersphere, *i.e.*, the surface $x^{\alpha'}x^{\alpha'} = r^2, \alpha' = 1',2',3'$, $r = \Re\sin\eta$. In the local 5-dimensional pseudo-spherical system, the curvature quantity of these circles can be obtained as[4]

$$B_a = \left\{\frac{1}{\Re},\frac{1}{\Re}\cot\eta,0,0,0\right\}, \quad a = 0,1,\cdots,4 \tag{2.14}$$

or with $r = \Re\sin\eta$, the more familiar form

$$B_a = \left\{\frac{1}{r}\sin\eta,\frac{1}{r}\cos\eta,0,0,0\right\}. \tag{2.15}$$

[3]The Christoffel symbols are not appropriate for this purpose.

[4]One can find a detailed calculation in [9].

The 0-dimension is the local extra dimension. Comparing the 4-dimensional part $\left\{\frac{1}{r}\cos\eta,0,0,0\right\}$ of this equation with (2.12), we find that the 4-dimensional space cannot be flat. Squaring (2.15), we get $B_aB^a=1/r^2$, r being the curvature radius of the greater circles of the spherical part of the model.

Performing a Lorentz transformation in the local [1,4]-slice, we get for (2.15)

$$B_a=\left\{\frac{1}{r}\sin\eta,\frac{1}{r}\cos\eta\cos i\chi,0,0,\frac{1}{r}\cos\eta\sin i\chi\right\} \tag{2.16}$$

where $i\chi$ is the Lorentz angle. In the case of transformation to a comoving system in a universe expanding in free fall, the relative motion of a comoving observer is geometrically defined and the relation

$$\cos\eta\cos i\chi=1 \tag{2.17}$$

is satisfied. It should be noted that the geometrical quantity $\cos\eta$ and the kinematical quantity $\cos i\chi$, the Lorentz factor of the motion, are closely related. Using this, we can easily derive

$$\sin i\chi=i\tan\eta\,.$$

Respecting these two relations, we can obtain from (2.16)

$$B_a=\left\{\frac{1}{r}\sin\eta,\frac{1}{r},0,0,i\frac{1}{r}\sin\eta\right\}. \tag{2.18}$$

Here, $\left\{\frac{1}{r},0,0\right\}$ is the spatial part of quantity B, which seems to be flat according to (2.12). Evidently, this is a consequence of Einstein's elevator principle that we have discussed in detail in paper [3]. However, all five components of (2.18) need to be considered. Again, the square of B is $1/r^2$, with r as the radii of the greater circles. As expected, the curvature of space turns out to be an invariant property. The same holds for quantity C mentioned in (2.12).

In addition, further slices of the pseudo-hypersphere are open pseudo circles

$$x^{0'^2}+x^{4'^2}=\Re^2\cos^2\eta$$

with radii $\Re\cos\eta$ and curvatures $1/\Re\cos\eta$, recalling $x^{4'}$ as an imaginary coordinate. This curvature is the cause of the force of acceleration in the dS cosmos. The latter is a component of the Ricci-rotation coefficients. It should be noted that for the transition of this quantity into a comoving system, the inhomogeneous transformation law of the Ricci-rotation coefficients is required. We have discussed this problem in the quoted papers.

Omitting the calculation of all the components of the Einstein tensor, we can circumvent the question of the curvature of the model. We could not find any controls in Melia's papers concerning the first three components of the Einstein tensor. They would exhibit the curvature radii of the normal and oblique slices of the pseudo-hypersphere, representing the positively curved universe.

xiii) As already mentioned in the earlier parts of this section, a transition from the dS model to the Subluminal Model is rather simple: the restriction

xiii) As already mentioned in the earlier parts of this section, a transition from the dS model to the Subluminal Model is rather simple: the restriction $\Re = const.$ needs to be expunged. Then $\Re$, the radius of the pseudo-hypersphere, *i.e.*, the radius of our universe, behaves as a function of time. The evolution of our universe can be described by a series of self-similar dS universes. Evidently, the metric of the dS universe does not contain any information regarding how the universe can develop or how to calculate the change in $\Re$. A second set of differential equations needs to be consulted. These are the contracted Bianchi identities $R_{[m\,n\cdot r\cdot\ \|s]}^{\ \ \ \ \ \ s} = 0$ that provide possible changes of the Riemann curvature tensor. To define a genuine expanding cosmological model, the following two differential equation systems are needed:

$$\begin{aligned} &\text{(I)} \quad R_{mn} - \frac{1}{2} g_{mn} R = -\kappa T_{mn} \\ &\text{(II)} \quad R_{m\ \|n}^{\ n} - \frac{1}{2} R_{\|m} = 0 \end{aligned} \quad . \tag{2.19}$$

System (II) leads to the conservation law $T_{m\ \|n}^{\ n} = 0$. This equation is often used in the literature to establish an outstanding relation with variables. Solving these two systems of equations, the Subluminal Model can be obtained with properties mentioned in items vii and viii.

3. Conclusion

We have demonstrated step by step that our positively curved and finite Subluminal Model is identical to Melia's $R_h = ct$ model and have extensively discussed the question of curvature. We conclude that Melia's claim that his model is flat is erroneous.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Burghardt, R. (2017) *Journal of Modern Physics*, **8**, 583-601. https://doi.org/10.4236/jmp.2017.84039

[2] Burghardt, R. (2019J) *Journal of Modern Physics*, **10**, 1439-1453. https://doi.org/10.4236/jmp.2019.1012096

[3] Burghardt, R. (2016) *Journal of Modern Physics*, **7**, 2447-2356. https://doi.org/10.4236/jmp.2016.716203

[4] Burghardt, R. (2018) *Journal of Modern Physics*, **9**, 685-700. https://doi.org/10.4236/jmp.2018.94047

[5] Florides, P.S. (1980) *General Relativity and Gravitation*, **12**, 563-574. https://doi.org/10.1007/BF00756530

[6] Burghardt, R. (2016) Transformations in de Sitter and Lanczos Models I. Austrian Reports on Gravitation. http://www.arg.or.at/Wpdf/WTrans1.pdf

[7] Burghardt, R. (2016) Transformations in de Sitter and Lanczos Models II. Austria

Reports on Gravitation. http://www.arg.or.at/Wpdf/WTrans2.pdf

[8] Burghardt, R. (2016) The Curvature Parameters in Gravitational Models. http://www.arg.or.at/Wpdf/Wcurv.pdf

[9] Burghardt, R. (2016) Spacetime Curvature. 1-597. http://arg.or.at/EMono.htm

Journal of Modern Physics, 2020, 11, 712-724
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Theoretical Prediction of Negative Energy Specific to the Electron

Koshun Suto

Chudai-Ji Temple, Isesaki, Japan
Email: Koshun_suto129@mbr.nifty.com

How to cite this paper: Suto, K. (2020) Theoretical Prediction of Negative Energy Specific to the Electron. *Journal of Modern Physics*, **11**, 712-724.
https://doi.org/10.4236/jmp.2020.115046

Received: April 9, 2020
Accepted: May 15, 2020
Published: May 18, 2020

Abstract

If an electron emits all of its rest mass energy $m_e c^2$, the relativistic energy of the electron will become zero. According to the special theory of relativity, an electron whose relativistic energy is zero does not have photon energy. In this paper, however, an electron is regarded as having photon energy $m_e c^2$ and negative energy $-m_e c^2$, even when its relativistic energy is zero. The state where relativistic energy is zero is achieved due to the positive energy and negative energy canceling each other out. Relativistic energy becomes zero for an electron in a hydrogen atom when the principle quantum number n is zero. The author has already pointed out the existence of an energy level with $n = 0$. If this model is used, it is possible for an electron in the state with $n = 0$ to emit additional photons, and transition to negative energy levels. The existence of negative energy specific to the electron has previously been nothing more than a conjecture. However, this paper aims to theoretically show the existence of negative energy based on a discussion using an ellipse. The results show that the electron has latent negative energy.

Keywords

Energy-Momentum Relationship in a Hydrogen Atom, Negative Energy Specific to the Electron, Relativistic Energy, Dark Matter

1. Introduction

The most important conclusion derived from the special theory of relativity (STR) is the equivalence of inertial mass and energy [1]. Energy in all its forms has inertial mass [2]. To put it another way, all changes in the energy of an object ΔE correspond to changes in the object's inertial mass Δm [3]. Einstein expressed these as follows.

$$E = mc^2. \tag{1}$$

$$\Delta E = c^2 \Delta m. \tag{2}$$

In this paper, let us review the energy-momentum relationship of Einstein using a textbook [4].

Now, in classical mechanics, an increase in kinetic energy corresponds to work done by an external force. That is,

$$\mathrm{d}E = F\mathrm{d}x = \frac{\mathrm{d}p}{\mathrm{d}t}\mathrm{d}x. \tag{3}$$

If $v = \mathrm{d}x/\mathrm{d}t$ is used here, Equation (3) can be written as follows.

$$\mathrm{d}E = v\mathrm{d}p. \tag{4}$$

The following relationship also holds in classical mechanics.

$$m = \frac{p}{v}. \tag{5}$$

In the textbook of French, the following equation is obtained by combining Equation (1) and Equation (5).

$$E = \frac{c^2 p}{v}. \tag{6}$$

Next, if the right-hand sides of Equation (4) and Equation (6), and the corresponding left hand sides, are multiplied together,

$$E\mathrm{d}E = c^2 p\mathrm{d}p. \tag{7}$$

Integrating this,

$$E^2 = c^2 p^2 + E_0^2. \tag{8}$$

Equation (8) can also be expressed as follows.

$$\left(mc^2\right)^2 = c^2 p^2 + \left(m_0 c^2\right)^2. \tag{9}$$

here m_0 is rest mass and m is relativistic mass. Equation (8) is the energy-momentum relationship of Einstein that holds in an isolated system in free space.

If an object is at rest, $p = 0$ and thus Equation (8) is as follows.

$$E = m_0 c^2. \tag{10}$$

However, if we are satisfied with Equation (10) only, then the deeper meaning of the theory of relativity is lost. Typically, momentum is not zero. In that case, Equation (1) is used.

Equation (1) includes Equation (10) as a special case. In this paper, the energy in Equation (1) becomes important when dealing with the relativistic energy of the hydrogen atom.

Now, what sort of relation holds in the case of an electron in a hydrogen atom?

Let's consider a situation where an electron at rest in free space is taken into a hydrogen atom due to the electrostatic attraction of the atomic nucleus (proton). At this time, the electron emits a photon and acquires kinetic energy. The in-

crease in kinetic energy of the electron corresponds to the work done with respect to the outside. This situation is the opposite of Equation (3), and thus Equation (3) must be rewritten as follows.

$$\mathrm{d}E = -F\mathrm{d}x = -\frac{\mathrm{d}p}{\mathrm{d}t}\mathrm{d}x. \tag{11}$$

From this, we obtain not Equation (4) but rather

$$\mathrm{d}E = -v\mathrm{d}p. \tag{12}$$

Multiplying in the same way the right-hand sides of Equation (6) and Equation (12), and the corresponding left hand sides,

$$E\mathrm{d}E = -c^2 p\mathrm{d}p. \tag{13}$$

And integrating,

$$E^2 = -c^2 p^2 + E_0^2. \tag{14}$$

Equation (14) can also be expressed as follows [5].

$$\left(m_n c^2\right)^2 = -c^2 p_n^2 + \left(m_\mathrm{e} c^2\right)^2, \quad m_n < m_\mathrm{e}. \tag{15}$$

Equation (15) is an energy-momentum relationship applicable to an electron in a hydrogen atom which has potential energy.

here, $m_\mathrm{e}c^2$ is the rest mass energy of the electron, and $m_n c^2$ is the relativistic energy of the electron. The subscript n is the principal quantum number.

Incidentally, the author has previously conjectured that the electron has latent negative energy in ref. [6]. However, since then no papers have been published developing the author's conjecture. Therefore, this paper aims to theoretically predict the existence of this negative energy.

2. Relativistic Correction of the Bohr Energy Levels

Past attempts to relativistically expand the energy levels of the hydrogen atom derived by Bohr have taken Equation (8) as their point of departure [7] [8] [9] [10]. However, this is a mistake.

The equation which treats an electron in a hydrogen atom relativistically is not the Dirac equation satisfying Equation (8). It must be another equation satisfying Equation (14). The author has already derived this equation [11].

The energy levels derived by Bohr are given by the following formulas [12] [13].

$$E_{\mathrm{BO},n} = -\frac{1}{2}\left(\frac{1}{4\pi\varepsilon_0}\right)^2 \frac{m_\mathrm{e}e^4}{\hbar^2}\cdot\frac{1}{n^2} \tag{16a}$$

$$E_{\mathrm{BO},n} = -\frac{\alpha^2 m_\mathrm{e}c^2}{2n^2}, \quad n = 1, 2, \cdots. \tag{16b}$$

here, $E_{\mathrm{BO},n}$ signifies the energy levels derived by Bohr. Also, α is the fine-structure constant, and is defined as follows.

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} = 7.2973525693\times10^{-3}. \tag{17}$$

In Bohr's theory, the energy when the electron has been pulled away from the atomic nucleus, and is stationary at an infinite distance, is set to be zero. However, the true energy in this state is $m_e c^2$. The relativistic energy $m_n c^2$ in Equation (15) indicates the energy with an absolute scale.

Incidentally, Einstein and Sommerfeld defined kinetic energy as the difference between the total energy and rest mass energy of an object [14]. That is,

$$K = mc^2 - m_0 c^2 = m_0 c^2 \left[\frac{1}{\left(1 - v^2/c^2\right)^{1/2}} - 1 \right]. \tag{18}$$

If this definition is also used for the kinetic energy of an electron in a hydrogen atom, then the energy levels of the hydrogen atom can be defined as follows.

$$E_{\mathrm{re},n} = -K_{\mathrm{re},n} = m_n c^2 - m_e c^2. \tag{19}$$

here, $K_{\mathrm{re},n}$ is the relativistic kinetic energy of an electron [15]. However, it is not appropriate to call $E_{\mathrm{re},n}$ relativistic energy. The relativistic energy of an electron has already been defined by Equation (15). $E_{\mathrm{re},n}$ gives energy levels which correct the Bohr energy levels (16) from the perspective of the theory of relativity. However, it should be noted that the term "theory of relativity" is applied not in the sense of Equation (9) but in the sense of Equation (15).

Incidentally, the following formula can be derived from Equation (9).

$$m = \frac{m_0}{\left(1 - v^2/c^2\right)^{1/2}}. \tag{20}$$

Using similar mathematics, the following formula can be derived from Equation (15) [16].

$$m_n = \frac{m_e}{\left(1 + v_n^2/c^2\right)^{1/2}}, \quad m_n < m_e. \tag{21}$$

An electron in a hydrogen atom becomes lighter in mass as it increases in speed. This is the opposite of the prediction of STR.

Next, Equation (21) is rewritten using the following relationship derived in ref. [15].

$$\frac{v_n}{c} = \frac{\alpha}{n}. \tag{22}$$

when that is done,

$$m_n = \frac{m_e}{\left(1 + \alpha^2/n^2\right)^{1/2}}. \tag{23}$$

Hence the energy levels in Equation (19) become as follows.

$$E_{\mathrm{re},n} = m_n c^2 - m_e c^2 = m_e c^2 \left[\left(\frac{n^2}{n^2 + \alpha^2} \right)^{1/2} - 1 \right], \quad n = 0, 1, 2, \cdots. \tag{24}$$

To simplify the discussion in this paper, the only quantum number addressed

is n.

Next, when the part of Equation (24) in parentheses is expressed as a Taylor expansion,

$$E_{\mathrm{re},n} \approx m_{\mathrm{e}}c^2\left[\left(1-\frac{\alpha^2}{2n^2}+\frac{3\alpha^4}{8n^4}-\frac{5\alpha^6}{16n^6}\right)-1\right] \tag{25a}$$

$$E_{\mathrm{re},n} \approx -\frac{\alpha^2 m_{\mathrm{e}}c^2}{2n^2}. \tag{25b}$$

It is evident from this that the Bohr energy levels (16b) are an approximation of Equation (24).

Incidentally, it was once pointed out by Dirac that Equation (8) has a negative solution [17]. In the same way, the author has pointed out that Equation (15) has a negative solution [18]. The mass of an electron at negative energy levels becomes negative.

In the current universe, there is thought to exist a tremendous mass whose true nature is unknown (an unknown source of gravity). The author has presented matter formed from an electron with negative mass and a proton (atomic nucleus) with positive mass as a strong candidate for this unknown matter, *i.e.*, dark matter [19] [20] [21] [22].

Now, if the negative solution of Equation (15) is also incorporated, then the relativistic energy $E^{\pm}_{\mathrm{ab\,re},n}$ of a hydrogen atom can be written as follows.

$$E^{\pm}_{\mathrm{ab\,re},n} = \pm m_n c^2 = \pm m_{\mathrm{e}}c^2 \mp K_{\mathrm{re},n} = \pm m_{\mathrm{e}}c^2\left(\frac{n^2}{n^2+\alpha^2}\right)^{1/2}, \quad n = 0,1,2,\cdots. \tag{26}$$

The $E_{\mathrm{ab\,re},n}$ newly introduced here has a long subscript. However, the reason for this is to distinguish from $E_{\mathrm{re},n}$ in Equation (19). $E_{\mathrm{ab\,re},n}$ is relativistic energy, but this is also energy expressed with an absolute scale. The "ab" in the subscript stands for "absolute," while "re" standard for "relativistic." If the above points are kept in mind, then there is no problem in abbreviating $E_{\mathrm{ab\,re},n}$ as $E_{\mathrm{ab},n}$.

3. Equation (15) Derived from an Ellipse

In this section, Equation (15) is derived using an ellipse.

First, consider the Cartesian coordinate system O-*xy*. Letting F and F' be the points $x = \pm f$, an ellipse is drawn taking those 2 points as foci (**Figure 1**).

Let A and A′ be the points where the ellipse intersects the *x*-axis, and let B and B′ be the points where the ellipse intersects the *y*-axis. Also, let $2a$ be the length of the line segment $\overline{\mathrm{AA'}}$, $2b$ be the length of the line segment $\overline{\mathrm{BB'}}$, and $2f$ be the length of the line segment $\overline{\mathrm{FF'}}$.

The eccentricity of the ellipse in this case is defined as follows.

$$\varepsilon = \frac{f}{a}. \tag{27}$$

The eccentricity of the ellipse can also be expressed using the following formula.

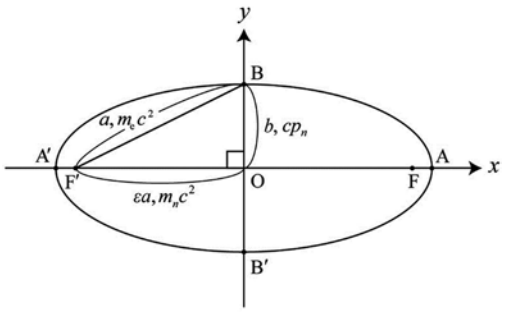


Figure 1. First, the energy $m_e c^2$ is taken to correspond to the line segment $\overline{\mathrm{OA}}$, and then Equation (31) is assumed. Equation (15) can be derived if the Pythagorean theorem is applied to the right triangle OBF'.

$$\varepsilon = \left(1 - \frac{b^2}{a^2}\right)^{1/2}. \tag{28}$$

The following equation can be derived from Equation (28).

$$b = a\left(1 - \varepsilon^2\right)^{1/2}. \tag{29}$$

here, the line segment $\overline{\mathrm{OA}}$ is taken to correspond to the energy $m_e c^2$. Let us express this as follows.

$$a = m_e c^2. \tag{30}$$

Also, assume the following relation.

$$\varepsilon = \left(\frac{n^2}{n^2 + \alpha^2}\right)^{1/2}. \tag{31}$$

Taking Equation (30) and Equation (31) into account, b can be expressed with the following equations.

$$b = m_e c^2 \left[1 - \left(\frac{n^2}{n^2 + \alpha^2}\right)\right]^{1/2} \tag{32a}$$

$$= m_e c^2 \left(\frac{\alpha^2}{n^2 + \alpha^2}\right)^{1/2}. \tag{32b}$$

Incidentally, if the length of the hypotenuse AB is taken to be c in the right triangle AOB in Figure 1, then the Pythagorean theorem below holds for the lengths of the three sides of the right triangle.

$$c^2 = a^2 + b^2. \tag{33}$$

Next, the Pythagorean theorem is applied to the right triangle OBF'. At that time, taking into account that $f = a\varepsilon$, we obtain

$$(\varepsilon a)^2 + b^2 = \left(\frac{n^2}{n^2 + \alpha^2}\right)\left(m_e c^2\right)^2 + \left(\frac{\alpha^2}{n^2 + \alpha^2}\right)\left(m_e c^2\right)^2 = \left(m_e c^2\right)^2. \tag{34}$$

here, if Equation (23) and Equation (22) are also taken into consideration, then the equation in Equation (34) can be expressed as follows.

$$\left(\frac{\alpha^2}{n^2+\alpha^2}\right)\left(m_e c^2\right)^2 = \left(\frac{\alpha^2}{n^2+\alpha^2}\right) m_n^2 \left(\frac{n^2+\alpha^2}{n^2}\right) c^2 \cdot c^2 = \frac{\alpha^2}{n^2} \cdot m_n^2 \frac{n^2 v_n^2}{\alpha^2} c^2 = c^2 p_n^2. \quad (35)$$

Substituting this result for Equation (35) into Equation (34),

$$\left(\varepsilon a\right)^2 + b^2 = \left(m_n c^2\right)^2 + c^2 p_n^2 = \left(m_e c^2\right)^2. \quad (36)$$

when $m_e c^2$ is taken to correspond to the line segment $\overline{\mathrm{OA}}$, and Equation (31) is assumed, then Equation (15) can be derived from the right triangle OBF'. Also, from Equation (30) and Equation (32b),

$$\frac{b}{a} = \left(\frac{\alpha^2}{n^2+\alpha^2}\right)^{1/2}. \quad (37)$$

when Equation (37) is substituted for α in Equation (17), it is evident that the ellipse that should actually be drawn is far flatter than that in **Figure 1**. However, to make the diagram easier to read, the ellipse is drawn with an elongated *y*-axis in this paper. (The segment $\overline{\mathrm{AF}}$ was also drawn longer than it actually is for the same reason.)

Incidentally, the author has previously derived the following relation in ref. [21].

$$\frac{r_n^-}{r_n^+} = \frac{\left(n^2+\alpha^2\right)^{1/2} - n}{\left(n^2+\alpha^2\right)^{1/2} + n} = \frac{m_e - m_n}{m_e + m_n}. \quad (38)$$

However, when considered using an ellipse, Equation (38) can be expanded as follows (**Figure 2**).

$$\frac{\overline{\mathrm{AF}}}{\overline{\mathrm{AF}'}} = \frac{a-f}{a+f} = \frac{1-\varepsilon}{1+\varepsilon} = \frac{r_n^-}{r_n^+} = \frac{\left(n^2+\alpha^2\right)^{1/2} - n}{\left(n^2+\alpha^2\right)^{1/2} + n} = \frac{\left(m_e - m_n\right)c^2}{\left(m_e + m_n\right)c^2}. \quad (39)$$

Also, the following relation can be obtained from **Figure 1**.

$$\frac{\overline{\mathrm{FF}'}}{\overline{\mathrm{AA}'}} = \varepsilon = \frac{\left(r_n^+ + r_n^-\right) - 2r_n^-}{r_n^+ + r_n^-} = \frac{r_n^+ - r_n^-}{r_n^+ + r_n^-} = \frac{m_n}{m_e} = \left(\frac{n^2}{n^2+\alpha^2}\right)^{1/2}. \quad (40)$$

In Equation (39), r_n^+ was taken to correspond to the line segment $\overline{\mathrm{AF}'}$, and r_n^- was taken to correspond to the line segment $\overline{\mathrm{AF}}$. Also, if $m_e c^2$ is taken to correspond to the line segment $\overline{\mathrm{OA}}$, based on Equation (40), then $2m_e c^2$ corresponds to line segment $\overline{\mathrm{AA}'}$. At this time, $\left(m_e - m_n\right)c^2$ corresponds to line segment $\overline{\mathrm{AF}}$, and $\left(m_e + m_n\right)c^2$ corresponds to the line segment $\overline{\mathrm{AF}'}$.

However, this does not mean that line segment length expresses the absolute value of a physical quantity, so caution is necessary. Also the orbital radii r_n^+ and r_n^- at different energy levels are apparent in the ellipse in **Figure 2**. This shows that this ellipse is not an illustration of the orbital of a single electron. The discussion thus far is summarized in the following table (**Table 1**).

here, r_e is the classical electron radius, defined as follows.

$$r_e = \frac{e^2}{4\pi\varepsilon_0 m_e c^2}. \quad (41)$$

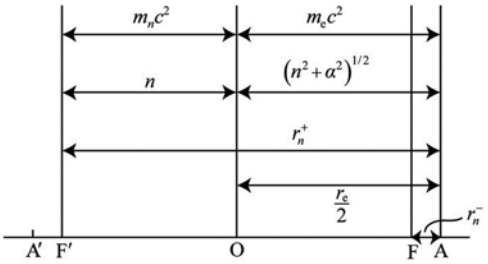


Figure 2. Caution is necessary because the lengths of the line segments in this diagram do not give the absolute value of physical quantities.

Table 1. Electron energy corresponding to each line segment and orbital radius. There are differences between energy *1 and *2, and energy *3 and *4. These energy differences are discussed in section 4.

Line Segment	$\overline{\mathrm{AF}}$	$\overline{\mathrm{AO}}$	$\overline{\mathrm{AF'}}$
Orbital Radius Corresponding to Line Segment	r_n^-	$\frac{r_e}{2}$	r_n^+
Relativistic Energy of Electron	$-m_n c^2$	0 *3	$m_n c^2$ *1
Indication using n and α	$\left(n^2+\alpha^2\right)^{1/2}-n$	$\left(n^2+\alpha^2\right)^{1/2}$	$\left(n^2+\alpha^2\right)^{1/2}+n$
Energy Corresponding to Line Segment	$\left(m_e-m_n\right)c^2$	$m_e c^2$ *4	$\left(m_e+m_n\right)c^2$ *2

Incidentally, when the electron's relativistic energy $m_n c^2$ becomes zero, then the value of $E_{\mathrm{re},n}$ in Equation (19) is

$$E_{\mathrm{re},0} = -m_e c^2, \quad m_n c^2 = 0. \tag{42}$$

Also, the following formula can be derived from the relationship between hydrogen atom energy levels and potential energy.

$$E_{\mathrm{re},n} = \frac{1}{2}V\left(r_n\right) = -\frac{1}{2}\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r_n} = -\frac{1}{2}\frac{r_e m_e c^2}{r_n} = -m_e c^2\left(\frac{r_e/2}{r_n}\right). \tag{43}$$

here, if Equation (42) and Equation (43) are compared, the radius r_0 when $n=0$, is

$$r_0 = \frac{r_e}{2}. \tag{44}$$

Incidentally, the electron does not have a definite orbital in quantum mechanics. The author has previously predicted in ref. [6] that the radius of the atomic nucleus of the hydrogen atom (proton) is $r_e/4$. However, this is not a claim that the atomic nucleus is completely spherical. The meaning is that statistically the atomic nucleus can be regarded as a sphere with radius $r_e/4$. In the same way, this paper does not assert that an electron with energy $E_{\mathrm{ab},n}^+$ revolves in an elliptical orbital with semi-major radius r_n^+. (In quantum mechanics, the Bohr radius a_B of the hydrogen atom signifies the radius where the probability of finding the electron is maximal.)

4. Discussion

1) In Figure 2 and Table 1, the relativistic energy corresponding to n is $m_n c^2$ (*1). The orbital radius r_n^+ of an electron with this energy becomes $\left(n^2+\alpha^2\right)^{1/2}+n$ when indicated in terms of n and α. However, the energy corresponding to this is $\left(m_\mathrm{e}+m_n\right)c^2$ (*2).

There are two types of electron energy in a certain state. How should we think about this? First of all, the following two points are correct.

a) The relativistic energy of an electron with orbital radius of r_n^+ is $m_n c^2$ (*1).

b) The energy of an electron with orbital radius of r_n^+ is $\left(m_\mathrm{e}+m_n\right)c^2$ (*2).

We must elucidate the true nature of the energy $\left(m_\mathrm{e}+m_n\right)c^2$. The author has previously assumed in ref. [6] that an electron in a state with $n=0$ has photon energy and negative energy. Let's use that as a hint.

Now, taking Figure 2 and the existence of negative energy into account, Table 1 can be expanded as follows (Table 2).

Here, if the photon energy possessed by an electron is expressed as $E_{\mathrm{ph},n}$, then the relativistic energy of the electron $E_{\mathrm{ab},n}$ defined in Equation (26) can be written as follows.

$$E_{\mathrm{ab},n}=E_{\mathrm{ph},n}-m_\mathrm{e}c^2. \tag{45}$$

According to this model, $m_n c^2$ (*1) is the relativistic energy of the electron, and $\left(m_\mathrm{e}+m_n\right)c^2$ (*2) is the photon energy possessed by the electron.

Incidentally, the relativistic energy of an electron $E_{\mathrm{ab},n}$ is thought to be what determines the electron's orbital radius. However, in Equation (45), the interpretation is also possible that it is $E_{\mathrm{ph},n}$ which determined the orbital radius of the electron.

Now, in the state where $E_{\mathrm{ab}}=0$, the photon energy possessed by the electron and the negative energy $-m_\mathrm{e}c^2$ cancel each other out (the relationship of *3, *4, and *5 in Table 2). If this model is used, an electron with $E_{\mathrm{ab}}=0$ can emit additional photons, and transition to negative energy levels.

Table 2. Relationship of the relativistic energy of an electron, and the photon energy possessed by that electron. The relativistic energy of an electron is defined as the sum of the photon energy possessed by that electron, and the negative energy specific to it.

Line Segment	—	$\overline{\mathrm{AF}}, a-f$	$\overline{\mathrm{AO}}, a$	$\overline{\mathrm{AF'}}, a+f$	—
Eccentricity Indication	—	$(1-\varepsilon)a$	a	$(1+\varepsilon)a$	—
Orbital Radius	$\frac{r_\mathrm{e}}{4}$	r_n^-	$\frac{r_\mathrm{e}}{2}$	r_n^+	∞
Indication using n and α	—	$\left(n^2+\alpha^2\right)^{1/2}-n$	$\left(n^2+\alpha^2\right)^{1/2}$	$\left(n^2+\alpha^2\right)^{1/2}+n$	—
Relativistic Energy of Electron $E_{\mathrm{ab},n}$, $E_{\mathrm{ph},n}-m_\mathrm{e}c^2$	$-m_\mathrm{e}c^2$	$-m_n c^2$	0 *3	$m_n c^2$ *1	$m_\mathrm{e}c^2$ *6
Photon Energy $E_{\mathrm{ph},n}$	0	$\left(m_\mathrm{e}-m_n\right)c^2$	$m_\mathrm{e}c^2$ *4	$\left(m_\mathrm{e}+m_n\right)c^2$ *2	$2m_\mathrm{e}c^2$ *7
Negative Energy $-m_\mathrm{e}c^2$	$-m_\mathrm{e}c^2$	$-m_\mathrm{e}c^2$	$-m_\mathrm{e}c^2$ *5	$-m_\mathrm{e}c^2$	$-m_\mathrm{e}c^2$ *8

Also, according to the STR, an electron has a rest mass energy of $m_e c^2$. However, in this paper, it is thought that an electron at rest has photonic energy $2m_e c^2$ and negative energy $-m_e c^2$ (the relationship of *6, *7, and *8 in **Table 2**). If symmetry is taken into account, it is likely valid to think that the true nature of this negative energy is a photon with negative mass. (However, that does not mean that in this paper it was possible to determine the amount of negative energy to be $-m_e c^2$.)

2) The author has previously derived the following formulas in ref. [21].

$$r_n^+ = \frac{r_e}{2}\frac{\left(n^2+\alpha^2\right)^{1/2}}{\left(n^2+\alpha^2\right)^{1/2}-n}. \tag{46}$$

$$r_n^- = \frac{r_e}{2}\frac{\left(n^2+\alpha^2\right)^{1/2}}{\left(n^2+\alpha^2\right)^{1/2}+n}. \tag{47}$$

Equation (46) and Equation (47) can be written as follows.

$$r_n^+ = \frac{r_e}{2}\left[1+\frac{n}{\left(n^2+\alpha^2\right)^{1/2}-n}\right]. \tag{48}$$

$$r_n^- = \frac{r_e}{2}\left[1-\frac{n}{\left(n^2+\alpha^2\right)^{1/2}+n}\right]. \tag{49}$$

Also, the following equation holds if Equation (39) is taken into account.

$$r_n^+ = \frac{r_e}{2}\cdot\frac{m_e}{m_e-m_n} = \frac{r_e}{2}\left(1+\frac{m_n}{m_e-m_n}\right). \tag{50}$$

$$r_n^- = \frac{r_e}{2}\cdot\frac{m_e}{m_e+m_n} = \frac{r_e}{2}\left(1-\frac{m_n}{m_e+m_n}\right),\quad \frac{r_e}{4} < r_e^-. \tag{51}$$

In addition, if the eccentricity of the ellipse is used,

$$r_n^+ = \frac{r_e}{2}\left[1+\frac{n}{\left(n^2+\alpha^2\right)^{1/2}-n}\right] = \frac{r_e}{2}\left(1+\frac{\varepsilon_n}{1-\varepsilon_n}\right). \tag{52}$$

$$r_n^- = \frac{r_e}{2}\left[1-\frac{n}{\left(n^2+\alpha^2\right)^{1/2}+n}\right] = \frac{r_e}{2}\left(1-\frac{\varepsilon_n}{1+\varepsilon_n}\right). \tag{53}$$

here, if Equation (48) and Equation (49), and Equation (50) and Equation (51) are respectively combined into single formulas, the results are as follows.

$$r_n^\pm = \frac{r_e}{2}\left[1\pm\frac{n}{\left(n^2+\alpha^2\right)^{1/2}\mp n}\right]. \tag{54}$$

$$r_n^\pm = \frac{r_e}{2}\left(1\pm\frac{m_n}{m_e\mp m_n}\right). \tag{55}$$

The orbital radius of an electron is defined as the distance from the center of the atomic nucleus to the electron. However, according to Equation (54) and Equation (55), the sizes of r_n^+ and r_n^- are determined taking $r = r_e/2$ as a starting point.

This is a discovery that has not been previously stated. r_n^- approaches the atomic nucleus (radius $r_e/4$) as n increases. In quantum mechanics, lower energy is regarded as more stable, but that is incorrect. Actually, the closer the relativistic energy of an electron is to zero, the greater the stability. Therefore, electrons with negative energy are never easily incorporated into the atomic nucleus.

Next, let us investigate the relationship between r_n^+ and the Bohr radius of the hydrogen atom. The Bohr radius is given by the following formula.

$$a_B n^2 = r_1 n^2 = 4\pi\varepsilon_0 \cdot \frac{\hbar^2}{m_e e^2} \cdot n^2 = \frac{r_e}{\alpha^2} \cdot n^2. \tag{56}$$

Next, Equation (46) is rewritten as follows.

$$r_n^+ = \frac{r_e}{2} \cdot \frac{\left(n^2+\alpha^2\right)^{1/2}}{\left(n^2+\alpha^2\right)^{1/2} - n} \cdot \frac{\left(n^2+\alpha^2\right)^{1/2} + n}{\left(n^2+\alpha^2\right)^{1/2} + n} \tag{57a}$$

$$= \frac{r_e}{2} \cdot \frac{n^2 + \alpha^2 + n^2\left(1+\alpha^2/n^2\right)^{1/2}}{\alpha^2}. \tag{57b}$$

when the part of the numerator in Equation (56b) that is in parentheses is developed as a Taylor expansion,

$$\left(1+\frac{\alpha^2}{n^2}\right)^{1/2} \approx 1 + \frac{\alpha^2}{2n^2} - \frac{\alpha^4}{8n^4} + \frac{\alpha^6}{16n^6}. \tag{58}$$

If this is substituted into Equation (57b),

$$r_n^+ = \frac{r_e}{\alpha^2} \cdot n^2 + \frac{r_e}{2}\left(\frac{3}{2} - \frac{\alpha^2}{8n^2}\right) \approx \frac{r_e}{\alpha^2} \cdot n^2 = a_B n^2. \tag{59}$$

From this, it is evident that the Bohr radius is an approximate value for Equation (46).

5. Conclusions

1) In this paper, the mathematically derived Equation (15) was also derived from a discussion using an ellipse. This paper was also able to theoretically predict the existence of latent negative energy possessed by the electron. The relationship is unclear between this negative energy, and the negative energy (dark energy) whose true nature is unknown, and which is currently thought to exist in large quantities in the universe. Further research will be needed in the future to elucidate that point.

2) The author has believed that the orbital radius of an electron forming a hydrogen atom is determined by the electron's relativistic energy $E_{ab,n}$. Even this is definitely not a mistake. However, this paper has shown more accurately that

what determines the orbital radius of the electron is the photon energy $E_{\mathrm{ph},n}$ possessed by the electron.

Also, the orbital radius of the electron is given as the distance from the center of the atomic nucleus (proton) to the electron. However, this paper has shown that the size of the electron orbital will be determined taking $r = r_{\mathrm{e}}/2$ as the starting point.

Acknowledgements

I would like to express my thanks to the staff at ACN Translation Services for their translation assistance. Also, I wish to express my gratitude to Mr. Shimada, H. for drawing figures.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Einstein, A. (1961) Relativity. Crown, New York, 43.

[2] Einstein, A. (1946) *Technical Journal*, **5**, 16.

[3] French, A.P. (1968) Special Relativity. W.W. Norton & Company, New York, London, 19.

[4] French, A.P. (1968) Special Relativity. W.W. Norton & Company, New York, London, 21.

[5] Suto, K. (2011) *Physics Essays*, **24**, 301-307. https://doi.org/10.4006/1.3583810

[6] Suto, K. (2018) *Applied Physics Research*, **10**, 93-101. https://doi.org/10.5539/apr.v10n4p93

[7] Kraft, D.W. (1974) *American Journal of Physics*, **42**, 837-839. https://doi.org/10.1119/1.1987875

[8] Cushing, J.T. (1970) *American Journal of Physics*, **38**, 1145. https://doi.org/10.1119/1.1976568

[9] Von Baeyer, H.C. (1975) *Physical Review D*, **12**, 3086. https://doi.org/10.1103/PhysRevD.12.3086

[10] Terzis, A.F. (2008) *European Journal of Physics*, **29**, 735-743. https://doi.org/10.1088/0143-0807/29/4/008

[11] Suto, K. (2018) *Applied Physics Research*, **10**, 102-108. https://doi.org/10.5539/apr.v10n6p102

[12] Bohr, N. (1913) *Philosophical Magazine*, **26**, 1. https://doi.org/10.1080/14786441308634955

[13] Bohr, N. (1952) Collected Works Vol. 2. North-Holland, Amsterdam, 136.

[14] Sommerfeld, A. (1923) Atomic Structure and Spectral Lines. Methuen & Co. Ltd., London, 528.

[15] Suto, K. (2019) *Applied Physics Research*, **11**, 19-34. https://doi.org/10.5539/apr.v11n1p19

[16] Suto, K. (2012) *Physics Essays*, **25**, 488-494.

https://doi.org/10.4006/0836-1398-25.4.488

[17] Dirac, P.A.M. (1978) Directions in Physics. Wiley, New York, 11. https://doi.org/10.1002/ev.1212

[18] Suto, K. (2014) *Applied Physics Research*, **6**, 64-73. https://doi.org/10.5539/apr.v6n6p64

[19] Suto, K. (2015) *Global Journal of Science Frontier Research A*, **15**, 1-6. https://journalofscience.org/index.php/GJSFR/article/view/1869

[20] Suto, K. (2017) *Applied Physics Research*, **9**, 70-76. https://doi.org/10.5539/apr.v9n1p70

[21] Suto, K. (2017) *Journal of Physical Mathematics*, **8**, 8.

[22] Suto, K. (2020) *Journal of High Energy Physics, Gravitation and Cosmology*, **6**, 52-61. https://doi.org/10.4236/jhepgc.2020.61007

Journal of Modern Physics, 2020, 11, 725-740
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

The Bell Inequalities: Identifying What Is Testable and What Is Not

Louis Sica[1,2]

[1]Institute for Quantum Studies, Chapman University, Orange, CA & Burtonsville, MD, USA
[2]Inspire Institute Inc., Alexandria, VA, USA
Email: lousica@edu.jhu

How to cite this paper: Sica, L. (2020) The Bell Inequalities: Identifying What Is Testable and What Is Not. *Journal of Modern Physics*, **11**, 725-740.
https://doi.org/10.4236/jmp.2020.115047

Received: April 15, 2020
Accepted: May 15, 2020
Published: May 18, 2020

Abstract

The Bell theorem and inequality were derived as consequences of seemingly reasonable physical and statistical hypotheses. Bell's assumptions were used to deduce cross-correlations of three spin measurements on two entangled particles neglecting non-commutation. The *assumed* correlation functions, later confirmed for certain quantum measurements, violate the Bell inequality. The present paper reviews a more general derivation of the Bell inequality showing that it is identically satisfied by finite data sets whether deterministic or random, after assuming merely that they exist. It is thereafter concerned with the consequences of this result for interpretations of the inequality that result in its violation. A primary finding is that correlation functions have differing forms due to quantum commutation, non-commutation, and conditions of measurement, and result in satisfaction of the Bell inequality used consistently with its derivation. A stochastic process having the same correlation function for all variable pairs is shown to be inconsistent with experimentally reported data. The logic of the three and four variable inequalities is shown to be similar. Finally the inequalities in probabilities are shown to follow from those in correlations with quantum mechanical results satisfying either when properly implemented.

Keywords

Bell Theorem, Bell Inequality, Entanglement, Locality, Correlations, Hidden Variables, Non-Commutation, Commutation, Cross-Correlations, Non-Stationary

1. Introduction

The purpose of this paper is to identify an oversight and accompanying errors in the logic of the Bell theorem. Violation of Bell inequalities by experimental data

results from misunderstanding the nature and processing of the data to be used. While quantum mechanics as a whole is not understood, and therefore admits various interpretations, the Bell inequality rests on mathematical logic alone. This has been unrecognized due to Bell's derivation, but is not a matter of interpretation once pointed out.

The Bell [1] inequality was originally derived as part of a theorem in statistics (See Appendix). However, the same inequality is derivable as a purely algebraic result that must be identically satisfied by three (or four) mutually cross-correlated data sets consisting of ±1's, regardless of whether they are random or deterministic [2]. The Bell inequality is thus independent of the statistical assumptions that Bell used and that have been assumed to be necessary to its derivation. Violation of the Bell inequality results from its misuse based on ignorance of its purely mathematical basis, and most surprisingly, from ignoring the established quantum mechanical principle of non-commutation.

When the inequality is applied to random processes, its basis in simultaneous cross-correlations limits the two-variable correlation functions that may occur in each triplet-of-variables or quadruplet-of-variables realization. From a mathematical perspective, there is no reason why the correlation functions among the different variable pairs need all be the same, and it will be found below that there is a simple reason why some should be qualitatively different.

Without even being aware of its name, the characteristics of a simple stochastic process known as wide sense stationary [3] (WSS) have been mistakenly assumed to characterize three measurements on two entangled particles involving non-commuting observables. That this error has not been recognized is possibly due to the stated belief that non-commutation is a purely quantum effect [4] that should not be considered in the context of possible hidden variables in quantum mechanics. In fact however, many classical processes are non-commutative and a major example of quantum non-commutation, the Pauli spin matrices, originated in a classical representation of three dimensional rotations by two dimensional matrices [5]. This, and the non-commutation of sequences of classical light polarization measurements, indicate that non-commutation is a fact that cannot be neglected in either classical or quantum physics. Indeed, while non-commutation is common in the classical world, the author is not aware of any books that treat it in the context of random variables until the recent monograph by Khrennikov [6]. Finally, although encountered frequently in everyday life, there is no commonly recognized English language term for non-commutation.

Bell based his theorem on the use of predicted quantum measurements (counterfactuals) [1] which some believe to be inapplicable to quantum mechanics. However, quantum mechanics may be used to provide probabilistic predictions for physical processes. Measurements may then be carried out to confirm predicted correlations [7]. As a result, the Bell inequalities may be applied to quantum counterfactuals that are subsequently measured as will be illustrated below.

A common explanation for the violation of the Bell inequality is that due to non-locality, more than three variables (or four) are actually interacting to produce the Bell cosine correlation. Then, the inequality is judged to be intrinsically inapplicable to the real physical situation. However, due to the extreme generality of the Bell inequality as derived below, simple procedures ensure the number of data sets necessary for inequality satisfaction, although alteration of the form of correlation functions may occur.

Finally, it is quite common to consider a probability version of the Bell inequality that some seem to assume is logically independent of the correlation form that Bell derived. However, since correlations may be expressed in terms of the probabilities that yield them, it is not surprising to find that the probability form follows from the Bell inequality in correlations. Thus the probability or Wigner [8] form of the Bell inequality is not logically independent of the correlational form and is satisfied by properly constructed quantum probabilities [9].

An earlier version of this paper was posted in the quant-ph archives [10].

2. Bell Inequality for Data Sets

The Bell inequality for data sets will now be derived, since it is of critical importance to understanding the Bell theorem. Data sets as defined herein, exist if they can be written down. They may result from experimental observations, from theoretical predictions of experimental observations (counterfactuals), or from a combination of the two. The data may also be either random or deterministic. In constructing his theorem and inequality, Bell assumed three measurement readouts from two entangled particles in a Bell-correlation measurement apparatus (**Figure 1**). There appear to be only two ways to obtain such data. One conceptually simple way is to add detectors (**Figure 2**) beyond the two needed to observe the correlated count pairs produced by the source. Such an arrangement immediately implies different correlations among different variable pairs.

Bell specifically rejected this approach [1]. He employed a second method using quantum probabilities to predict the result of an additional measurement at an alternative detector setting for a previously measured particle. However, only one measurement on each particle is commutative. In quantum mechanics, if two operators corresponding to measurement operations commute, they have a set of eigenstates or eigenfunctions in common. If they do not commute, their eigenfunctions are from different function sets with members of either set expressible as a linear combination of those of the other (see basic quantum mechanics texts, e.g., Mandl [11]). Thus, the probabilites for predicted correlations resulting from non-commuting alternative observables at different instrument settings should be different, and they are as will be shown below (See [7] for a suggested expermental test of this).

The Bell inequality will now be shown to hold under far more general conditions than is apparent from Bell's derivation as part of a statistics theorem [2].

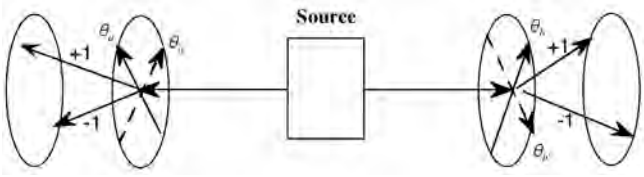


Figure 1. Schematic of Bell experiment in which a source sends two particles to two detectors having angular settings θ_a and θ_b and/or counterfactual settings $\theta_{a'}$ and $\theta_{b'}$. While one measurement operation on the A-side, e.g. at setting θ_a, commutes with one on the B-side at θ_b, any additional measurements at either $\theta_{a'}$ or $\theta_{b'}$ are non-commutative with prior measurements at θ_a and θ_b, respectively.

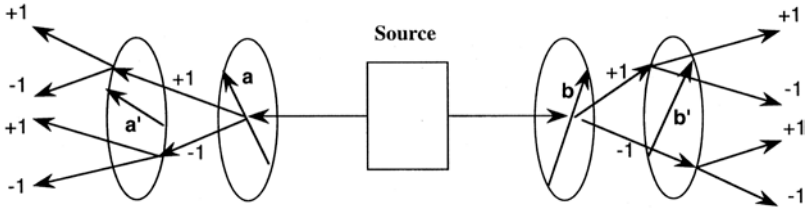


Figure 2. Schematic of a Multiple Stern-Gerlach apparatus. Arrows a, a', b, b' indicate the magnetic field directions encountered by pairs of particles emitted in opposite directions by the source. At each encounter with a magnetic field, the particle is deflected in one of two directions depending on whether its spin is +1/2 or −1/2. Each sequence of ±1's corresponds to a unique output position so that knowledge of two spin measurements is yielded retrospectively for each particle.

Assume that three data sets, random or deterministic, labeled a, b, and c have been obtained. The data set items are denoted by a_i, b_i, and c_i with N items in each set. Each datum equals ±1. One may form the equation

$$a_i b_i - a_i c_i = a_i \left(b_i - c_i \right), \tag{2.1}$$

and sum this equation over the N data triplets from the data sets. After dividing by N, one obtains

$$\frac{\sum_i^N a_i b_i}{N} - \frac{\sum_i^N a_i c_i}{N} = \frac{\sum_i^N a_i b_i \left(1 - b_i c_i \right)}{N}. \tag{2.2}$$

Taking absolute values of both sides,

$$\left| \frac{\sum_i^N a_i b_i}{N} - \frac{\sum_i^N a_i c_i}{N} \right| = \left| \frac{\sum_i^N a_i b_i \left(1 - b_i c_i \right)}{N} \right| \le \frac{\sum_i^N \left| \left(1 - b_i c_i \right) \right|}{N}, \tag{2.3}$$

or

$$\left| \frac{\sum_i^N a_i b_i}{N} - \frac{\sum_i^N a_i c_i}{N} \right| \le 1 - \frac{\sum_i^N b_i c_i}{N}. \tag{2.4}$$

Inequality (2.4) has the same form as the three variable inequality derived by Bell for correlations but is expressed in a form directly applicable to laboratory data. The algebraic steps used in it's construction are the same as those applied by Bell to previously averaged correlation functions (See Appendix).

The author unexpectedly discovered this result some time ago [2] by asking

the following question: If one performs a laboratory experiment for which the number of data items 3N is necessarily finite, to what extent do random fluctuations of the correlation estimates result in violation of the Bell inequality (2.4)? Surprisingly, the answer as shown above, is that the inequality is precisely satisfied. No assumption has been made other than that the data can be written down. Further, as long as the data can be tagged with labels a, b, and c, the inequality is satisfied even if nonlocl pickup exists between detectors. In that event, the form of the correlations would be affected but not whether the inequalitiy is satisfied. Indeed, in an extreme case, the data averaged correlations might not converge to identifiable limits, but the Bell inequality would still be satisfied for any three identified data sets. No experimental loophole in this conclusion is apparent. The inequality still follows if some of the ±1's are replaced by zeros.

If, however, the data derive from a random process, and the correlation estimates in inequality (2.4) converge to probabilistically computable correlations as $N \to \infty$, the resulting correlations designated by $C(x,y)$ must then satisfy an inequality of the same form as inequality (2.4):

$$\left|C(a,b)-C(a,c)\right| \le 1-C(b,c). \tag{2.5a}$$

where $C(a,b)=\lim_{N\to\infty}\sum_i \frac{a_i b_i}{N}$,

and the limit is statistical.

This is essentially the inequality derived by Bell [1] using a stochastic process model (see Appendix) in which detectors at the same settings on opposite sides of a source of entangled particle pairs produce results of opposite sign (**Figure 1**). In Bell's stochastic process model, left-hand-side detectors of **Figure 1** were designated $A(a,\lambda)$ and right-hand-side detectors $B(b,\lambda)$ with $A(a,\lambda)=-B(a,\lambda)$ so as to automatically agree with results of entanglement at equal settings. In deriving the right-hand side correlation, Bell represented the product of both outputs using one stochastic-model function, $A(b,\lambda)A(c,\lambda)$. However, the final correlation on the right-hand side is preceded by a plus sign that arises due to measurements being taken on opposite sides of the apparatus.

In some discussions below, measurement settings will be labeled to agree with the side on which the measurement occurs. That usage is consistent with Bell's stated interpretation of the variables in the three variable inequality ([1] Chapter 8) as predicted results: a measurement at setting *a*, followed by two *alternative* predictions of measurement results at *b* and *b*′. This will be treated in detail in Sec. 4.

In the optical case two polarizations occur. Counts of one polarization are labeled +1 and those of the orthogonal polarization are labeled −1. Using b' for the alternative measurement instead of *c* on the right-hand-side, the inequality is written

$$\left|C_1(a,b)-C_2(a,b')\right| \le 1-C_3(b,b'). \tag{2.5b}$$

Now, all a-measurements occur on one side of an apparatus such as shown in Figure 1 and b-measurements on the other.

The subscripts on the correlations in inequality (2.5b) indicate that *they do not all necessarily have the same functional form* as follows from the lack of conditions used in deriving inequality (2.4). This is directly relevant to the quantum mechanics (QM) case to which inequality (2.5b) is applied, and for which the correlations are different as results from non–commutation of measurements.

It is obvious that the constant equal to 1 occurring in inequality (2.5b) results from the fact that the value a_i that multiplies b_i also multiplies b_i', data triplet by data triplet. However, if data are obtained from an independent run for each correlated measurement pair as is common in practice, then six data sets instead of three are used, the condition under which the inequality was derived does not hold, and the inequality will in general be violated. Strangely, the use of independent runs has become accepted experimental practice. It is critically important to understand that while inequality (2.4) holds generally for any three arbitrary data sets, the Bell inequalities (2.5a, b) do not hold for arbitrary correlations. Since correlations must result from the convergence of the correlation estimates that satisfy inequality (2.4), they must satisfy inequalities (2.5a, b) and their functional forms are mutually constrained thereby. Arbitrary correlations not derivable as limiting forms of correlations of three concurrently existing data sets will not necessarily satisfy inequalities (2.5a, b). If inequalities (2.5a, b) are violated by assumed limiting forms, no data sets of triplets can exist that produce them.

Similar assumptions to those used to derive inequality (2.5a) from inequality (2.4) can also be used to derive a four variable Bell inequality. Assuming that there exist four data sets of size N with members a_i, a_i', b_i, b_i' with each datum equal to ±1, then for each group of four data items from the four data sets, one has (by inspection)

$$-2 \le a_i\left(b_i + b_i'\right) + a_i'\left(b_i - b_i'\right) \le 2\,. \tag{2.6}$$

(Inequality (2.6) also holds if zeros occur among the variables.) Summing over i from 1 to N in inequality (2.6), and dividing by N leads to

$$-2 \le \frac{\sum_i^N a_i b_i}{N} + \frac{\sum_i^N a_i b_i'}{N} + \frac{\sum_i^N a_i' b_i}{N} - \frac{\sum_i^N a_i' b_i'}{N} \le 2\,. \tag{2.7a}$$

Since all experimental data sets are intrinsically finite, four data sets must satisfy inequality (2.7a), as three must satisfy inequality (2.4). Again assuming statistical convergence to limits as $N \to \infty$, a common form of Bell inequality used by experimentalists results:

$$-2 \le C_1(a,b) + C_2(a,b') + C_3(a',b) - C_4(a',b') \le 2\,. \tag{2.7b}$$

As in the case of inequality (2.5b), the correlations may have different functional forms.

The difficulty of applying the three variable inequality (2.4) to an entangled

pair of particles in which more than two measurements are non-commutative is amplified in the case of a four variable inequality. Note, as in the previous case, that while inequality (2.7a) must be identically satisfied by any four data sets that may be written down, inequality (2.7b) may be violated by assumed correlations. However, as in the three variables case, if it is violated it follows that no four data sets exist whose cross-correlations result in the assumed correlation functions.

Note that the notion of experimentally "testing" the above inequalities, in either three or four variables, involves a logical mischaracterization. *Only the form of the several correlations that describe data from a given physical experiment may be tested, and not whether or not cross-correlations of the data satisfy the Bell inequalities*. (In the laboratory counts are observed, and correlations computed from them.) Further, if variables are obtained from random realizations that yield measurements on only one pair per realization among the four correlated variables, the correlations will in general be different than if all four variable values are obtained per realization. A conceptually simple way to obtain four data outputs per realization is shown in **Figure 2**.

3. How Independent Data-Pair Correlations May Violate Bell Inequalities

3.1. Bell Operationally Assumed Correlations That Are Wide-Sense Stationary

Inequalities in three and four variables were derived from Bell's assumption of a stochastic process representation of quantum entanglement [1]. Bell represented detector readouts with a function $A(a,\lambda)=\pm 1$, where a is an instrument setting and λ denotes one or more random variables determining the resulting random values taken by $A(a,\lambda)$. In Bell's representation, there is no implication that accessing a readout at a affects the probability of accessing a readout at a' for a given realization. The multiple readouts of function $A(a,\lambda)$ and their associated probabilities are analogous to a set of commuting observables in QM. Howeever, for the case of non-commuting QM observables that applies here, probabilities of specific readouts at successive instrument settings are conditionally dependent on readouts at prior settings [11]. The conditional probabilites at alternative instrument settings have different values, whereas they would be expected to be the same if commutative states were involved.

Without stating it, Bell effectively and operationally assumed properties for the several quantum mechanical outputs with which he was concerned that correspond to a special, not universal, kind of random process defined as wide-sense-stationary [3] (WSS). Such a process is one in which the correlation of the readouts $A(a,\lambda)$ at any two instrument settings a_i and a_j is given by a function of the form $f(a_i-a_j)$ depending on the difference of coordinate settings for all setting pairs. Thus, in Bell's notation

$$f(a_i-a_j)=\int_{-\infty}^{\infty}A(a_i,\lambda)A(a_j,\lambda)\rho(\lambda)\mathrm{d}\lambda \tag{3.1}$$

where $\rho(\lambda)$ is a probability distribution for λ, and a_i and a_j are any detector settings.

Bell used the correlation functional form computed from QM for commuting measurements on a pair of entangled spins that suggests a WSS process, $-\cos(a-b)$. The measurements commute because they are carred out on two different particles and it does not matter which measurement occurs first. Bell computed QM correlations at a setting *a* for a first detector and two alternative settings *b* and b' for a second detector ([1], Chap 8). Predicted QM correlations $C(a,b)$ and $C(a,b')$ are both given by the negative cosine of detector angular differences (suggesting WSS). However, if the resulting correlation $C(b,b')$ at output settings *b* and b' is computed from the different QM probabilities that occur for each non-commuting variable and fixed value *a* (that Bell assumed), the result is $C(a,b)C(a,b')$ as will be shown below.

3.2. How Misinterpretation of the Bell Inequality Leads to Its Violation

Inequalities (2.5b) and (2.7b) result from the cross-correlations of three and four data sets, respectively. A triplet or quadruplet of data values must occur in one realization of the associated random process. Whether three or four variables, each equal to ±1 are cross-correlated, determines the constant of the related inequality, 1 in the case of three variables, 2 in the case of four variables. However, if correlations are obtained from three or four variable *pairs*, with each pair acquired in an independent experimental run, the correlations will in general be quantitatively different from those that result from cross-correlated data triplets or quadruplets acquired in one run (as could be accomplished with the setup of **Figure 2**). Different measurement scenarios for the random variables will in general affect both correlations between variables and their corresponding probabilities.

Given measurement apparatus such as shown in **Figure 1**, two possible scenarios are identified for acquisition of three correlations. One may measure output pairs in independent runs at settings (a,b), (a,b'), and (b',b), (the third pair of settings on opposite sides of the apparatus) producing separate realizations of each variable pair. In that case, the correlation of each variable pair is given by the same function in the quantum situation under consideration, but the conditions that lead to Bell's derivation of the inequality as well as inequality (2.4), are violated.

A second scenario (that specified by Bell) is to predict the three outputs at settings *a*, *b*, and b' (with *b* and b' both on the right-hand side) for each random realization and calculate resulting correlations $C(a,b)$, $C(a,b')$, and $C(b',b)$ from QM probabilities for the variable pairs. Clearly the correlations and probabilities should be different from those obtained in the first scenario. Given that probabilities $P_{x,y}(a,b)$, $(x,y\in\pm1)$ are known from QM for measurements at (a,b) and (a,b') given setting *a*, one can immediately compute $P(b|a)$ allowing the evaluation of $P(b,b'|a)=P(b|a)P(b'|a)$. Thus, the now con-

nected correlations of outputs at setting pairs (a,b), (a,b'), and (b,b') may be determined. Hess has pointed out [12] that similar facts and inequalities related to those of Bell have been known to mathematicians since Boole. Pure mathematics determines a third correlation when data for two correlations out of the three are specified.

3.3. WSS Correlations Satisfy the Bell Inequality But They Are Not Co-Sinusoidal

It is instructive to consider inequalities (2.7a, b) for a finite value of N in the special case of a WSS process and four data sets. The WSS properties have been assumed to represent entanglement by experimentalists and theoreticians alike, after Bell's mistaken assumption of their universal applicability. One may write inequality (2.7a) in the form

$$\begin{aligned} -2 \le C(\theta_a - \theta_b) + \delta_N(a,b) + C(\theta_a - \theta_{b'}) + \delta_N(a,b') \\ + C(\theta_{a'} - \theta_b) + \delta_N(a',b) - C(\theta_{a'} - \theta_{b'}) - \delta_N(a',b') \le 2 \end{aligned} \tag{3.2}$$

where the $C(x-y)$ functions are assumed to represent the limiting forms for the correlation estimates as $N \to \infty$. Since the inequality cannot be violated by data sets that are jointly present and cross-correlated, the δ_N's represent random differences from the probability averaged correlation that lead to inequality satisfaction when the four variables' values are present in each realization of the experiment.

By contrast if the data are taken in four independent runs using the same instrument settings, inequality (3.2) for the same WSS process becomes

$$\begin{aligned} -2? \le C(\theta_a - \theta_b) + \delta_N(a_1,b_1) + C(\theta_a - \theta_{b'}) + \delta_N(a_2,b_2') \\ + C(\theta_{a'} - \theta_b) + \delta_N(a_3',b_3) - C(\theta_{a'} - \theta_{b'}) - \delta_N(a_4',b_4') \le 2? \end{aligned} \tag{3.3}$$

where the subscripts 1 … 4 indicate the experimental run number used to compute the correlation statistical fluctuation, and the question marks indicate possible violation of the ±2 limits since the data are no longer cross-correlated.

Note: the *cross correlation* of the data sets used in inequality (2.7a) is what causes that inequality to be identically satisfied and have a limiting magnitude of 2. If eight data sets and not four are used, the δ_N's plus the limiting correlations need no longer satisfy the inequality, even though the limiting correlations are given by the same function for the WSS process assumed. Given that the estimates statistically converge, the δ_N's are expected to become small as N becomes lager. Thus, although the inequality (3.3) would be violated, it would be violated by smaller and smaller values as N increases.

3.4. Quantum Mechanical Bell Correlations Cannot Represent a WSS Process

Bell effectively assumed [1] that the random process applicable to a triplet of polarization or spin measurement correlations is WSS, as is also widely done in the four variable case of inequality (2.7b) by those interpreting experimental data in

a way that violates inequality (2.7b). When the mathematical facts leading to inequalities (2.5) or (2.7b) are considered, however, it becomes clear that the measurement results in QM experiments do not represent a WSS process. If they did, violation of the corresponding Bell inequality would be expected to be small, *i.e.*, of the order of four standard deviations, rather than 102 standard deviations as has been reported [13]. Such inequality violations represent proof that the correlations of the process under consideration cannot all be co-sinusoidal and indeed they are not, as will be shown below. In the case of QM entanglement, only the measurements that commute between two particles are of this form.

4. The Wigner Inequality Results from a Bell Inequality If Probabilities Are Symhmetric

A probability inequality known as the Wigner [8] inequality is intimately related to the Bell inequality constraint on correlations [9]. It relates the probabilites for pairs of +1 outcomes corresponding to Bell correlations at given instrument settings. The result is

$$P_{++}(a,b) \le P_{++}(a,c) + P_{++}(a' = b,c). \tag{4.1}$$

where the first letter in each probability $P_{++}(a,b)$ indicates the angular setting for the random variable on the left side of the Bell apparatus in **Figure 1**, and the second letter indicates the angular setting for the random variable on the right side. The subscripts + and – indicate whether the variables at settings a and b have values of +1 or −1. Inequality. (4.1) follows from the Bell inequality

$$\left|C(a,b) - C(a,c)\right| \le 1 - C(b,c). \tag{4.2a}$$

If setting b of the right-hand-side correlation specifies instead the a-side setting of a Bell apparatus, the same numerical output occurs with a reversed sign under Bell's random process model where $A(b,\lambda) = -B(b,\lambda)$. Then inequality (4.2a) becomes

$$\left|C(a,b) - C(a,c)\right| \le 1 + C(b,c). \tag{4.2b}$$

A physical process is now considered with probabilities having the same symmetry of occurence for ±1's as occurs in QM in the case of two entangled spins. The joint probabilites resulting from an entangled spin state are:

$$\begin{aligned} P(b = \pm 1, a = \pm 1) &= \frac{1}{2}\sin^2\left(\frac{\theta_b - \theta_a}{2}\right); \\ P(b = \pm 1, a = \mp 1) &= \frac{1}{2}\cos^2\left(\frac{\theta_b - \theta_a}{2}\right). \end{aligned} \tag{4.3}$$

where a contracted notation indicates possible values of random variables b and a at settings θ_b and θ_a. The probabilities conditional on a are easily obtained since $P(a = \pm 1) = 1/2$. The joint probabilities (4.3) thus have the following symmetry for ±1 occurrence:

$$P_{++}(a,b) = P_{--}(a,b), \; P_{+-}(a,b) = P_{-+}(a,b). \tag{4.4}$$

The normalization condition is

$$2P_{++}(a,b)+2P_{+-}(a,b)=1. \tag{4.5a}$$

so that $C(a,b)$ is given by

$$C(a,b)=2P_{++}(a,b)-2P_{+-}(a,b). \tag{4.5b}$$

or

$$C(a,b)=4P_{++}(a,b)-1 \tag{4.5c}$$

after using Equation (4.5a). The results for $C(a,c)$ and $C(b,c)$ are obtained by renaming the variables in Equation (4.5c). The use of correlation (4.5c) in inequality (4.2b) with appropriate variables for different correlations produces:

$$4P_{++}(a,b)-1-\left(4P_{++}(a,c)-1\right)\le 1+4P(b,c)-1,$$

or the Wigner inequality:

$$P_{++}(a,b)\le P_{++}(b,c)+P_{++}(a,c) \tag{4.6}$$

As is well known, inequality (4.6) is violated when the same quantum Bell-state probability is used for all terms.

To show that quantum mechanical probabilities are consistent with a probability form of the Bell inequality it is simpler to use inequality (2.5b) since that form directly represents the physical situation considered by Bell [1]. In Chapter 8 of Bell's book, Speakable and unspeakable in quantum mechanics, Bell indicates that the result for the variable at b' is a predicted value on the same side of the apparatus as b.

The computation of $C(b,b')$ in terms of $P_{++}(b,b')$ (where both b-variables are now on the B-side of the apparatus) follows if the symmetry of Equations (4.4) holds for the probabilities of the observed and predicted variables:

$$\begin{aligned} C(b,b') &= \sum_a \left[P_{++}(b,b'\,|\,a)+P_{--}(b,b'\,|\,a)\right]P(a) \\ &\quad -\sum_a \left[P_{-+}(b,b'\,|\,a)+P_{+-}(b,b'\,|\,a)\right]P(a). \end{aligned} \tag{4.7a}$$

The joint probability $P_{++}(b,b')$ computed from probabilities (4.3) is

$$\begin{aligned} &P_{++}(b,b') \\ &=\frac{1}{2}\left[P_{+}(b\,|\,a=1)P_{+}(b'\,|\,a=1)+P_{+}(b\,|\,a=-1)P_{+}(b'\,|\,a=-1)\right], \\ &=\frac{1}{2}\left[\sin^2\frac{\theta_b-\theta_a}{2}\sin^2\frac{\theta_{b'}-\theta_a}{2}+\cos^2\frac{\theta_b-\theta_a}{2}\cos^2\frac{\theta_{b'}-\theta_a}{2}\right]. \end{aligned} \tag{4.7b}$$

A similar calculation for $P_{--}(b,b')$ yields the same result. Similarly, $P_{+-}(b,b')=P_{-+}(b,b')$ with result

$$P_{+-}(b,b')=\frac{1}{2}\left[\sin^2\frac{\theta_b-\theta_a}{2}\cos^2\frac{\theta_{b'}-\theta_a}{2}+\cos^2\frac{\theta_b-\theta_a}{2}\sin^2\frac{\theta_{b'}-\theta_a}{2}\right] \tag{4.7c}$$

Since the probabilities leading to $C(b,b')$ have the same symmetries as those for $C(a,b)$ and $C(a,b')$, and the probabilities for variable pairs are equal, they are used to compute $C(b,b')$ from Equation (4.7a). However, although the specified symmetries are the same, the probabilities are very different from probabilities (4.3) that lead to the Bell correlation for the first measurement pair.

5. Quantum Counterfactual Probabilities and Correlations Satisfy the Bell Inequality

5.1. Satisfying the Bell Inequality

Inequality (4.6) is violated by quantum probabilities upon assuming that all corresponding correlations have the same form as those for the two commuting measurements. This occurs because the correlation on the right-hand-side of inequality (2.5a) is constrained by the left-hand correlations $C(a,b)$ and $C(a,c)$ whose existence requires data that determines the right-hand side. (Note: $(ab)(ac)=bc$.) The correlation $C(b,c)$ thereby determined cannot have the same form as the previous correlations if the latter have the Bell cosine form. Repeating inequality (2.5b):

$$\left|C(a,b)-C(a,b')\right|\le 1-C(b,b'). \tag{5.1a}$$

Using the probabilities for $P_{++}(b,b')$ and $P_{+-}(b,b')$ given above in Equations (4.7b) and (4.7c) the resulting correlation $C(b,b')$ may be computed as [7]

$$\begin{aligned} C(b,b') &= 2P_{++}(b,b')-2P_{+-}(b,b') \\ &= \left[\sin^2\frac{\theta_b-\theta_a}{2}\sin^2\frac{\theta_{b'}-\theta_a}{2}+\cos^2\frac{\theta_b-\theta_a}{2}\cos^2\frac{\theta_{b'}-\theta_a}{2}\right] \\ &\quad -\left[\sin^2\frac{\theta_b-\theta_a}{2}\cos^2\frac{\theta_{b'}-\theta_a}{2}+\cos^2\frac{\theta_b-\theta_a}{2}\sin^2\frac{\theta_{b'}-\theta_a}{2}\right] \\ &= \cos(\theta_b-\theta_a)\cos(\theta_{b'}-\theta_a). \end{aligned} \tag{5.1c}$$

The same result may also be computed by using conditional probabilites for data outputs at b and b' given outputs for a. This result could be tested experimentally as suggested in Ref [7] where an analogous result is also given in the four variable case. Using the contracted notation $\theta(b,a)\equiv\theta_b-\theta_a$, and the Bell cosine correlations, inequality (5.1a) becomes

$$\begin{aligned} &\left|-\cos\theta(a,b)+\cos\theta(a,b')\right| \\ &\le 1-\cos\theta(b,a)\cos\theta(b',a)=1-\cos\theta(a,b)\cos\theta(a,b'), \end{aligned} \tag{5.1d}$$

from which

$$\begin{aligned} &\left|-2\sin\frac{\theta(a,b)+\theta(a,b')}{2}\sin\frac{\theta(a,b')-\theta(a,b)}{2}\right| \\ &\le \sin^2\frac{\theta(a,b)+\theta(a,b')}{2}+\sin^2\frac{\theta(a,b)-\theta(a,b')}{2}, \end{aligned} \tag{5.1e}$$

after use of appropriate trigonometric identities. One may replace the difference of correlations on the left-hand-side of inequality (5.1d) by expressions in probabilities P_{++} but the same result occurs in inequality (5.1e). Since

$$\begin{aligned} &\left(a^2+b^2\right)-2|a||b|=\left(|a|-|b|\right)^2\ge 0, \\ &\left(a^2+b^2\right)\ge 2|a||b|, \end{aligned} \tag{5.2}$$

the Bell inequality (5.1a) is satisfied.

Thus, when correlations computed from probabilities resulting from QM are used, whether expressed in correlational or probability form, inequality (5.1a) is satisfied as demanded by basic mathematics. Deductions of non-locality or non-reality, if based on Bell inequality violation, no longer follow.

5.2. Dealing with Possible Pickup between Detectors

If measurements are made on two particles, one of the measurements occurs first, except in circumstances of infinite time precision. Assuming that A is measured before B or B' by some time increment, any assumed pickup from other detectors by A is fixed when A is measured. If there is also pickup from detector A to B or A to B', three data sets are still obtained. Thus the three variable inequality holds even for the corrupted data. Observed correlation functions could then be compared with QM predictions to determine if evidence of pickup in fact exists. The Bell inequality would still be satisfied by the data sets even if the correlation estimates failed to converge to identifiable functions.

6. Conclusions

The principle claim with which this article is concerned is that correlations of quantum mechanical laboratory data violate the Bell inequality. Since it has been shown that the same inequality holds identically for data sets as a fact of algebra without Bell's assumptions, the notion that it is testable rests on a mathematical oversight. This has resulted in misuse of an inequality that must be identically satisfied when used correctly. What may be experimentally tested is not whether the Bell inequality is satisfied when correctly used, but the form of the multiple correlation functions realized from qualitatively different measurements. If correctly computed, correlations consistent with quantum mechanics do not all have the cosine form that Bell and others have assumed based on independent count-pair measurements. The Bell inequality constants result from data triplets and quadruplets obtained per realization, and not data pairs.

Understanding of the Bell inequality follows simply in the absence of logical errors. The three and four variable inequalities are identically satisfied by cross-correlations of finite quantum mechanical data sets of ±1's as a fact of algebra. Their satisfaction by quantum measurements follows from a well-known quantum mechanical fact: performed measurements on spins or photons are non-commutative when more than one per particle of an entangled pair is invoked. When both facts are employed, the Bell inequality is satisfied without mysteries.

The Bell theorem has been interpreted to imply that one cannot construct a local probability model that accounts for quantum correlations without entanglement. If the logic of the theorem is flawed, however, it does not follow thereby that the converse is true. The question that immediately arises is: how much of conventional quantum mechanics is to be accounted for. Since a local algorith-

mic model for Bell correlations has been presented [14], the elimination of non-locality would seem to be an attainable goal. This is in agreement with the observation that the physical superposition of four waves that produce entanglement in a down-converter source no longer exists when the waves propagate to spatially separated detectors for count detection. In this case one may derive a local probability model resulting from the boundary conditions that exist at the source [15]. This model appears to be consistent with both quantum electro-dynamics and wave optics.

The literature relevant to Bell's theorem has grown to a large size. There are growing numbers of papers that disagree with the Bell consensus according to [16]. The author apologizes in advance to those not cited that agree with him, and to those not agreeing as well. However, this already quite long article is concerned with the logical components of this controversial topic, rather than a review of the literature. A more uniform coverage of recent contributions will have to wait until a later date.

Acknowledgements

The author is indebdted to Joe Foreman for many conversations that motivated writing the present paper and to Michael Steiner for careful specification of some of the reasoning that has been applied to Bell inequality violation. Useful conversations with Armen Gulian have also aided the writing of this paper.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Bell, J.S. (1987) Speakable and Unspeakable in Quantum Mechanics. Cambridge University Press, Cambridge.

[2] Sica, L. (1999) *Optics Communications*, **170**, 55-60. https://doi.org/10.1016/S0030-4018(99)00417-4

[3] Papoulis, A. and Pillai, S.U. (2002) Probability, Random Variables, and Stochastic Processes. McGraw-Hill Companies, Inc., New York, Chap. 9.

[4] Griffiths, R.B. (2019) *Physics Today*, **72**, 11-13. https://doi.org/10.1063/PT.3.4306

[5] Goldstein, H. (1980) Classical Mechanics. Addison Wesley Publishing Co., Inc., Boston, 156.

[6] Khrennikov, A. (2016) Probability and Randomness Quantum versus Classical. Imperial College Press, London. https://doi.org/10.1142/p1036

[7] Sica, L. (2002) *Foundations of Physics Letters*, **15**, 473-486. https://doi.org/10.1023/A:1023920230595

[8] Wigner, E.P. (1970) *American Journal of Physics*, **38**, 1005-1009. https://doi.org/10.1119/1.1976526

[9] Sica, L. (2004) *Journal of Modern Optics*, **51**, 2461-2468. https://doi.org/10.1080/09500340408231804

[10] Sica, L. (2019) The Bell and Probability Inequalities Are Not Violated When Non-Commutation Is Applied According to Quantum Principles. https://export.arxiv.org/ftp/arxiv/papers/1903/1903.02116.pdf

[11] Mandl, F. (1957) Quantum Mechanics. Butterworths Scientific Publications, London, 75.

[12] Hess, K. (2015) Einstein Was Right. Pan Stanford Publishing Pte. Ltd., Singapore, Chap. 8.

[13] Kwiat, P.G., *et al.* (1995) *Physical Review Letters*, **75**, 4337-4342. https://doi.org/10.1103/PhysRevLett.75.4337

[14] de Raedt, H., Michielsen, K. and Hess, K. (2017) *Open Physics*, **15**, 713-733. https://doi.org/10.1515/phys-2017-0085

[15] Sica, L. (2019) Bell Correlations from Previously Entangled States of Light. Quant-ph arXiv:1908.06196

[16] de Raedt, H., Hess, K. and Michielsen, K. (2011) *Journal of Computational and Theoretical Nanoscience*, **8**, 1011-1039. https://doi.org/10.1166/jctn.2011.1781

Appendix

It is useful to derive Bell's version of his inequality by applying an explicit probability average to the left hand side of Equation (2.2):

$$\begin{aligned} &Av\left(\frac{1}{N}\sum_i^N a_i b_i - \frac{1}{N}\sum_i^N a_i b_i'\right) \\ &= \frac{1}{N}\int \sum_i^N a_i(\lambda)\left(b_i(\lambda) - b_i'(\lambda)\right)\rho(\lambda)\mathrm{d}\lambda. \end{aligned} \tag{A1}$$

In (A1), $\rho(\lambda)$ is a probability density for a collection of random variables λ that determine the values of the data items a_i, b_i, and b_i'. Changing to Bell's random process notation for which $A(a,\lambda) = \pm 1$ at setting *a*, $B(b,\lambda) = \pm 1$ at setting *b*, and $B(a,\lambda) = -A(a,\lambda)$, each function determined by random variable λ, Equation (A1) becomes

$$\begin{aligned} &(1/N)\int \sum_i^N a_i(\lambda)\left(b_i(\lambda) - b_i'(\lambda)\right)\rho(\lambda)\mathrm{d}\lambda \\ &= \int A(a,\lambda)\left(B(b,\lambda) - B(b',\lambda)\right)\rho(\lambda)\mathrm{d}\lambda, \end{aligned} \tag{A2}$$

where the probability averages are independent of subscript *i* since they are the same for each *i*. The three variables A, B, B' have random values determined by the collection of random parameters λ, and define a stochastic process. The right side of (A2) then becomes

$$\begin{aligned} &\int A(a,\lambda)\left(B(b,\lambda) - B'(b',\lambda)\right)\rho(\lambda)\mathrm{d}\lambda \\ &= \int A(a,\lambda)B(b,\lambda)\left(1 - B(b,\lambda)B'(b',\lambda)\right)\rho(\lambda)\mathrm{d}\lambda. \end{aligned} \tag{A3}$$

Taking absolute values of both sides and bringing the absolute value inside the integral on the right leads to

$$\begin{aligned} &\left|C(a,b) - C(a,b')\right| \\ &\le \int \left|A(a,\lambda)B(b,\lambda)\right|\left|\left(1 - B(b,\lambda)B'(b',\lambda)\right)\right|\rho(\lambda)\mathrm{d}\lambda \\ &= \int \left|\left(1 - B(b,\lambda)B'(b',\lambda)\right)\right|\rho(\lambda)\mathrm{d}\lambda \\ &= \int \left(1 - B(b,\lambda)B'(b',\lambda)\right)\rho(\lambda)\mathrm{d}\lambda \\ &= 1 - C(bb') \end{aligned} \tag{A4}$$

where $C(a,b)$ indicates the probability average of $A(a,\lambda)B(b,\lambda)$ etc. Relations (A4) end in Bell's inequality [1].

Bell's notation suggests that any number of readouts may be obtained for a given realization λ of the process. It is consistent with a WSS process as described above. The interpretation used thus specifies a particularly simple random process that is by no means universal. The result (2.4) proves that the same relation holds independently of Bell's stated assumptions in proving version (A4), and even if the data are deterministic.

Journal of Modern Physics, 2020, 11, 741-765
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Proton and Neutron Electromagnetic Form Factors Based on Bound System in 3 + 1 Dimensional QCD

Teruo Kurai

1-32-1 Nishifu-machi, Fuchu-shi, Tokyo, Japan
Email: kurai.teruo@topaz.plala.or.jp

How to cite this paper: Kurai, T. (2020) Proton and Neutron Electromagnetic Form Factors Based on Bound System in 3 + 1 Dimensional QCD. *Journal of Modern Physics*, **11**, 741-765.
https://doi.org/10.4236/jmp.2020.115048

Received: March 27, 2020
Accepted: May 18, 2020
Published: May 21, 2020

Abstract

We propose a new description of a nucleon as a pair of pions. The baryon number of our description of nucleon is not 1 but 0. However, this is probable because the proton spin crisis shows that the baryon spin cannot tell the number of composing quarks anymore. Because we use the derived pion wave function to describe a nucleon, our description has automatically the pionic degree of freedom and can be compared to the constituent quark model. Using this description, we investigate the electric charge and magnetization density functions of protons and neutrons. The electric charge density function of neutron is quite similar to those of Galster model and Maints data except the magnitude of singularity. The density functions of proton also show the similar behavior as those of Kelly's except near origin. Taking the Fourier transform of the density functions, we obtain the Sachs electromagnetic form factors that can be compared to those in the parametrization derived by Ye *et al.*

Keywords

Density Function, Electromagnetic Form Factor

1. Introduction

The proton and neutron electromagnetic form factors (e.m. FFs) are key components for understanding the charge and magnetization distributions within nucleons. In the past 20 years, a new generation of experiments, frequently utilizing polarization of freedom, has provided new knowledge regarding our understanding of the form factors [1] [2]. The parametrization work proposed by Ye *et al.* [3] to analyze the form factors including uncertainties used the com-

plete world data set for electron scattering and applied their best knowledge of two photon exchange (TPE) corrections. From the viewpoint of the charge and magnetization distributions within nucleons, which are principal reasons to investigate e.m. FFs, we need to investigate the wave functions (WFs) of nucleon directly. Under the constituent quark model (CQM), Chung and Coester [4] developed the light-front calculation of the nucleon FF using a Gaussian WF in the quark internal (transverse) momentum variables. This model yields good agreement with the observed $G_E^{[P]}/G_M^{[P]}$ ratio, but its nucleon FF decreases too quickly at larger $\left|\vec{Q}\right|^2$. Schlumpf [5] used the power law dependence of the quadratic internal momentum variables in the nucleon light-front WF and showed reasonable results for the power behavior of the FF at larger $\left|\vec{Q}\right|^2$. The WF of Schlumpf was used by Frank [6] and Millar [7]. Cardarell [8] used the rest-frame WF obtained by the Capstich-Isgur model [9] and found it to yield a significant content of high-momentum components, which are generated by the short range portion of quark-quark interactions. A comparable amount of high-momentum components in the nucleon WF was obtained in the Goldstone-boson-exchange quark model (GBE CQM), and this led to the nucleon e.m. FFs in point-like form [10] [11] [12] [13] [14]. The covariant CQM calculation yielding fairly good agreement with the nucleon FF was performed by Gross and Agbite [15]. For the pion cloud model, Miller [16] performed a light-front cloudy bag model calculation. This chiral quark model includes the perturbative pions, and is improved by Faessler [17]. A non-perturbative approach that combines both quark and pion degrees of freedom and interpolates between CQM and the SKyrum model (where the nucleon appears as a soliton solution of an effective nonlinear pion field) is the chiral soliton model (χQSM), in which the baryon density is not exactly Gaussian but is quite close to it. Another approach to the estimation of Sachs e.m. FFs results from the generalized parton distribution scheme (GPD) [18] [19] [20]. In this scheme, the three-parameter modified Regge mode provides a good description in the range of low $\left|\vec{Q}\right|^2$ to large $\left|\vec{Q}\right|^2$, but the ratio of $G_E^{[P]}/G_M^{[P]}$ becomes negative beyond 8 (GeV^2), which does not fit the data [21]. Because this analysis was conducted in momentum space, it was necessary to take the Fourier Transform to obtain the charge and magnetization density functions (distribution functions). In this paper, we propose a new description of a nucleon, *i.e.*, a pion pair, in configuration space (r-space) and show that the charge and magnetization density functions follow directly from this description. Taking the Fourier transform, we derive the Sachs e.m. FFs, which are comparable to those in Ye *et al.* [3]. This means that we can investigate the Sachs e.m. FFs and the charge and magnetization density functions with opposite ordering, in contrast to Kelly's way [22].

To clarify, we list here the symbols and parameters.

$\rho_{ch}^{[P]}, \rho_m^{[P]}$: electric charge and magnetization density functions of proton.

$\rho_{ch}^{[N]}, \rho_m^{[N]}$: electric charge and magnetization density functions of neutron.

$\tilde{\rho}_{ch}^{[P]}, \tilde{\rho}_{m}^{[P]}$: proton intrinsic FFs that are Fourier transform of electric charge and magnetization density functions.

$\tilde{\rho}_{ch}^{[N]}, \tilde{\rho}_{m}^{[N]}$: neutron intrinsic FFs that are Fourier transform of electric charge and magnetization density functions.

n_P, n_N: parameters of the power of exponentials of proton and neutron these are used for both density functions and Sachs e.m. FFs.

m_P, m_N: parameters of the coefficient of r or $\left|\vec{Q}\right|^2$.

β: range parameter for density functions.

p: parameter of the coefficient of logarithmic term of Sachs e.m. FFs.

2. Derivation

2.1. Basic Concept

To date, there have been several proposed descriptions for a nucleon. One of these is the pion cloud model, which introduces an elementary, perturbative pion couple to the constituent quark model (CQM) such that chiral symmetry is restored. Noting the fact that the contribution of quark spins to the spin of a proton is small, *i.e.*, the proton spin crisis [23], it is a fair consideration that nucleons are described only by the pion pair as far as the e.m. FFs are concerned. Baryon number of our case is not 1 but 0, however, proton crisis tells that the composing quark intrinsic spin does not contribute to that of baryon. This indicates that the baryon spin cannot tell the number of the composing quarks. The reason is following. It is quite reasonable consideration that the spin of all baryons are not determined by their composing quark intrinsic spin because all baryons decay to proton as a final state. The baryon number 1 comes from the assumption that the number of quarks is three for all baryons, and this assumption is based on the consideration that the composing quark intrinsic spin determines that of baryon. Proton spin crisis shows that this assumption lost the experimental support. Then, it is probable to consider that baryon number is not 1. Thus, we propose the description of nucleon as a pion pair even though baryon number is 0. Normally, the simplest description of a proton is uud (up, up, down) and that of a neutron is udd (up, down, down), while our description of proton is a $\pi^+ - \pi^0$ pair composed of $u\bar{d} - \left(u\bar{u} + d\bar{d}\right)/\sqrt{2}$, and that of a neutron is a $\pi^+ - \pi^-$ pair composed of $u\bar{d} - \bar{u}d$. The most important feature of the WFs is to describe the e.m. FFs because the electric charge and the magnetization density functions (distribution functions) are directly connected to them, as shown by Kelly [22]. We derived the pion $\left(\pi^+\left(\pi^-\right)\right)$ WF based on the hadronic operator proposed by Suura [24] [25] in a previous paper [26]. The derived charged pion WF that we use in this paper is Gaussian/r in configuration space (r space) thus the density function is $\text{Gaussian} \times F\left(1, \frac{3}{2}; Q^2\right)$ in momentum space as shown later section. $F\left(1, \frac{3}{2}; Q^2\right)$ is the one of Kummer's confluent hyper geometric series. Therefore this is quite similar to the constituent

quark model (CQM) using with Schlumpf's Q^2 polynomials momenta. Therefore we can expect our e.m. FFs become similar to these of CQM except their behavior in larger Q^2 region because two pion moves independently in this region as we explain later section. Thus, as far as e.m. FFs are concerned, we consider a proton and a neutron to be described as $\pi^+ - \pi^0$ and $\pi^+ - \pi^-$, respectively.

2.2. Basic Concept of Evaluation

The Bethe-Salpeter-like amplitude of the hadronic operator applied in this paper is defined as

$$\chi(1,2) = \langle 0 | q(1,2) | P \rangle \tag{1}$$

where $|0\rangle$ and $|P\rangle$ denote a vacuum and the physical state, respectively.

The gauge-invariant bi-local operator $q(1,2)$ is defined in the non-Abelian gauge field as

$$q(1,2) = T_r^c q_\beta^+(2) P \exp\left(ig \int_1^2 \mathrm{d}\vec{x} \vec{A}^a(x) \frac{\lambda_a}{2} \right) q_\alpha(1) \tag{2}$$

here α and β denote the Dirac indices, P denotes the path ordering, and the $\frac{\lambda_a}{2}$ components are generators of the adjoint representation of the SU (N). color gauge group. The trace is calculated for color spin a. Suura first proposed this definition [24] [25], and later applied it to the case of the light meson mass spectra, the t'Hooft model, and the pion e.m. FF [26] [27] [28]. In the case of pions, quarks (and antiquarks) are specified as u ($\bar{u}$) and d ($\bar{d}$), however, as we previously described [26], we obtained the pion wave function with eigenvalues and eigen functions of the equation of motion for a $\chi(1,2)$ system. Thus, we can describe the $\pi^+ - \pi^0$ pair as $\chi(1,2)$ and $\chi(3,4)$. Here, we describe π^0 as just one Bethe-Salpeter-like amplitude even though π^0 is desdribed as $\left(u\bar{u} + d\bar{d}\right)/\sqrt{2}$. As mentioned in section 2, the derived pion WF in Ref. [26] is for a charged pion. We cannot expect the WF of π^0 to be the same as that of charged pion, but the mass of π^0 is very close to that of charged pion. Thus, we assume that π^0 is also described as the same form of the Bethe-Salpeter-like amplitude of the hadronic operator as that of the charged pion.

2.3. Evaluation

The basic concept of our evaluation is as follows.

Because baryons are represented as initially binding meson pairs as described in subsection 5, at $\left|\vec{Q}\right|^2 = 0$ and for the small $\left|\vec{Q}\right|^2$ case, the $\pi^+ - \pi^0$ pair has the same origin. For simplicity, both of quarks are in the same position. However, when $\left|\vec{Q}\right|^2$ is large, both π^+ and π^0 gradually move freely with respect to each other and both π^+ and π^0 move totally independent of each other,

which is as the same concept as asymptotic freedom.

2.3.1. Density Functions

$$\begin{aligned} q(1;2,3) = \frac{1}{2}\Bigg[& \left(T_r^c q^\dagger(2) P \exp\left(ig\int_1^2 \mathrm{d}\vec{x}\,\vec{A}^a(x)\frac{\lambda_a}{2}\right) q(1)\right)^{(\pi^+)} \\ & + \left(T_r^c q^\dagger(3) P \exp\left(ig\int_1^3 \mathrm{d}\vec{x}\,\vec{A}^a(x)\frac{\lambda_a}{2}\right) q(1)\right)^{(\pi^0)} \Bigg] \\ & + \frac{1}{2}\Bigg[\left(T_r^c q^\dagger(2) P \exp\left(ig\int_1^2 \mathrm{d}\vec{x}\,\vec{A}^a(x)\frac{\lambda_a}{2}\right) q(1)\right)^{(\pi^0)} \\ & + \left(T_r^c q^\dagger(3) P \exp\left(ig\int_1^3 \mathrm{d}\vec{x}\,\vec{A}^a(x)\frac{\lambda_a}{2}\right) q(1)\right)^{(\pi^+)} \Bigg] \end{aligned} \tag{3}$$

At first ($\left|\vec{Q}\right|^2 = 0$), the $\pi^+ - \pi^0$ pair is described as where $(\pi^+), (\pi^0)$ means that each described hadronic operator corresponds to π^+ and π^0, respectively.

Note that Equation (3) promises the evenly charged wave function and that the total factor $\frac{1}{2}$ keeps proton charge at $+e$, and that the positions of both π^+ and π^0 quarks are at the same point ($\left|\vec{Q}\right|^2 = 0$).

The latter $\frac{1}{2}$ part is essentially same as the former $\frac{1}{2}$ part when we do not consider the charge distribution. Because our concern is the wave function of the $\pi^+ - \pi^0$ system, the Bethe-Salpeter-like amplitude can be defined as

$$\chi(1;2.3) = \langle 0 | q(1;2.3) | P \rangle \tag{4}$$

where

$$q(1;2,3) = q(1,2) + q(1,3)$$

here, we dropped the factor $\frac{1}{2}$ for simplicity. When considering charge distribution function, we consider this factor again.

Then equation of motion of $q(1;2,3)$ becomes

$$\begin{aligned} i\frac{\partial}{\partial t} q(1;2,3) &= i\frac{\partial}{\partial t} q(1,2) + i\frac{\partial}{\partial t} q(1,3) \\ &= -i\vec{\alpha}\cdot\vec{\nabla}(2) q(1,2) - q(1,2) i\vec{\alpha}\cdot\vec{\nabla}(1) + g\int_1^2 \mathrm{d}\vec{x} q_{\vec{E}}(1,2;x) \\ &\quad + g\vec{\alpha}\cdot\int_1^2 \mathrm{d}\vec{x}\times q_{\vec{B}}(1,2;x) - i\vec{\alpha}\cdot\vec{\nabla}(3) q(1,3) - q(1,3) i\vec{\alpha}\cdot\vec{\nabla}(1) \\ &\quad + g\int_1^3 \mathrm{d}\vec{x} q_{\vec{E}}(1,3;x) + g\vec{\alpha}\cdot\int_1^3 \mathrm{d}\vec{x}\times q_{\vec{B}}(1,3;x) \\ q_{\vec{\sigma}}(1,s;x) &= q^+(s) U(s,x) \vec{O}^a \frac{\lambda_a}{2} U(x,1) \end{aligned} \tag{5}$$

where

$$U(s,1) \equiv P\exp\left(ig\int_1^s \mathrm{d}\vec{x}\vec{A}^a(x)\frac{\lambda_a}{2} \right)$$

$$\alpha^k = \gamma^0\gamma^k$$

O is any operator, $s = 2,3$ and γ^μ is γ is matrices.

We previously obtained the equation of $i\frac{\partial}{\partial t}q(1,s)$ [28].

Because we consider every quark mass to be zero, the center of mass coordinate and two relative coordinates can be written

$$\begin{aligned} \vec{G} &= \frac{\vec{r}_1 + \vec{r}_2 + \vec{r}_3}{3} \\ \vec{r}^{(1)} &= \vec{r}_2 - \vec{r}_1,\ \vec{r}^{(2)} = \vec{r}_3 - \vec{r}_1 \end{aligned} \tag{6}$$

Then

$$\begin{aligned} \vec{r}_1 &= \vec{G} - \frac{1}{3}\vec{r}^{(1)} - \frac{1}{3}\vec{r}^{(2)} \\ \vec{r}_2 &= \vec{G} + \frac{2}{3}\vec{r}^{(1)} - \frac{1}{3}\vec{r}^{(2)} \\ \vec{r}_3 &= \vec{G} - \frac{1}{3}\vec{r}^{(1)} + \frac{2}{3}\vec{r}^{(2)} \end{aligned} \tag{7}$$

Thus, the derivatives are

$$\begin{aligned} \vec{\nabla}(1) &= \frac{\partial}{\partial\vec{r}_1} = \frac{1}{3}\frac{\partial}{\partial\vec{G}} - \frac{\partial}{\partial\vec{r}^{(1)}} - \frac{\partial}{\partial\vec{r}^{(2)}} \\ \vec{\nabla}(2) &= \frac{1}{3}\frac{\partial}{\partial\vec{G}} + \frac{\partial}{\partial\vec{r}^{(1)}} \\ \vec{\nabla}(3) &= \frac{1}{3}\frac{\partial}{\partial\vec{G}} + \frac{\partial}{\partial\vec{r}^{(2)}} \end{aligned} \tag{8}$$

Remembering that $\vec{\alpha}\cdot\vec{\nabla}(s) = \alpha^l\frac{\partial}{\partial r_s^l}$ ($s = 1,2,3$), the kinetic term becomes

$$\begin{aligned} &-i\vec{\alpha}\cdot\vec{\nabla}(2)q(1,2) - q(1,2)i\vec{\alpha}\cdot\vec{\nabla}(1) \\ &= -\frac{i}{3}\left\{ \alpha^l\frac{\partial}{\partial G^l}, q(1,2) \right\}_+ - i\left[\alpha^l\frac{\partial}{\partial r^{(1)l}}, q(1,2) \right] + q(1,2)i\alpha^l\frac{\partial}{\partial r^{(2)l}} \end{aligned} \tag{9}$$

Similarly

$$\begin{aligned} &-i\vec{\alpha}\cdot\vec{\nabla}(3)q(1,3) - q(1,3)i\vec{\alpha}\cdot\vec{\nabla}(1) \\ &= -\frac{i}{3}\left\{ \alpha^l\frac{\partial}{\partial G^l}, q(1,3) \right\}_+ - i\left[\alpha^l\frac{\partial}{\partial r^{(2)l}}, q(1,3) \right] + q(1,3)i\alpha^l\frac{\partial}{\partial r^{(1)l}} \end{aligned} \tag{10}$$

We consider the gauge field string only for the straight line case. Thus, the hadronic operator $q(1,2)$ is decomposed in the relative coordinate system as

$$q(1,2) = q_0\left(r^{(1)}\right) + i\vec{\alpha}\cdot\hat{r}^{(1)}q_1\left(r^{(1)}\right) + \beta q_2\left(r^{(1)}\right) + \beta\left(i\vec{\alpha}\cdot\hat{r}^{(1)}\right)q_3\left(r^{(1)}\right) \tag{11}$$

where $r^{(1)} = \left|\vec{r}^{(1)}\right|$

Because $q_0\left(r^{(1)}\right) = 0, q_1\left(r^{(1)}\right) = 0$ (as we previously showed [26]), only term

we have to deal with is $i\frac{\partial}{\partial r^{(2)l}}\left(\beta\left(i\vec{\alpha}\cdot\hat{r}^{(1)}\right)q_3\left(r^{(1)}\right)\right)$.

$$-\beta i\alpha^l\frac{\partial}{\partial r^{(2)l}}i\alpha^m\frac{x^{(1)m}}{r^{(1)}}=\beta\alpha^l\alpha^m\frac{\partial}{\partial x_l^{(2)}}\left(\frac{x_m^{(1)}}{r^{(1)}}\right)=0 \tag{12}$$

$$\frac{\partial}{\partial x^{(2)l}}q_s\left(r^{(1)}\right)=0\ \ (s=2,3) \tag{13}$$

here, we denote that $r^{(p)l}=x_l^{(p)}$ ($p=1,2$), so that $r^{(1)}=\sqrt{x_1^{(1)2}+x_2^{(1)2}+x_3^{(1)2}}$

Thus,

$$q(1,2)i\alpha^l\frac{\partial}{\partial r^{(2)l}}=0 \tag{14}$$

Similarly

$$q(1,3)i\alpha^l\frac{\partial}{\partial r^{(1)l}}=0 \tag{15}$$

Therefore, the kinetic terms in the relative coordinate system become

$$\begin{aligned}&-i\vec{\alpha}\cdot\vec{\nabla}(2)q(1,2)-q(1,2)i\vec{\alpha}\cdot\vec{\nabla}(1)-i\vec{\alpha}\cdot\vec{\nabla}(3)q(1,3)-q(1,3)i\vec{\alpha}\cdot\vec{\nabla}(1)\\&=-i\left[\alpha^l\frac{\partial}{\partial r^{(1)l}},q(1,2)\right]-i\left[\alpha^l\frac{\partial}{\partial r^{(2)l}}\right]\end{aligned} \tag{16}$$

The integral terms in relative coordinate system become

$$g\int_1^2\mathrm{d}\vec{x}q_{\vec{E}}(1,2;x)=-\frac{g^2}{2}\int_0^{r^{(1)}}\mathrm{d}zq\left(r^{(1)}-z\right)\left(r^{(1)}-z\right)q(z) \tag{17}$$

$$g\vec{\alpha}\cdot\int_1^2\mathrm{d}\vec{x}\times q_{\vec{B}}(1,2;x)=\frac{g^2}{2}\int_{-\infty}^t\mathrm{d}t'\left(\vec{\alpha}\cdot\hat{r}^{(1)}\right)\delta(t-t')\int_0^{r^{(1)}}\mathrm{d}zq\left(t',r^{(1)}-z\right)q(t',z) \tag{18}$$

$$g\int_1^3\mathrm{d}\vec{x}q_{\vec{E}}(1,3;x)=-\frac{g^2}{2}\int_0^{r^{(2)}}\mathrm{d}zq\left(r^{(2)}-z\right)\left(r^{(2)}-z\right)q(z) \tag{19}$$

$$g\vec{\alpha}\cdot\int_1^3\mathrm{d}\vec{x}\times q_{\vec{B}}(1,3;x)=\frac{g^2}{2}\int_{-\infty}^t\mathrm{d}t'\left(\vec{\alpha}\cdot\hat{r}^{(2)}\right)\delta(t-t')\int_0^{r^{(2)}}\mathrm{d}zq\left(t',r^{(2)}-z\right)q(t',z) \tag{20}$$

We obtained these equations were obtained previously [26].

Thus, the equation of motion for $q(1;2,3)$ is expressed by the following independent equations in the relative coordinate system.

$$\begin{aligned}i\frac{\partial}{\partial t}q\left(r^{(1)}\right)=&-i\left[\vec{\alpha}\cdot\vec{\nabla}_{r^{(1)}},q\left(r^{(1)}\right)\right]-\frac{g^2}{2}\int_0^{r^{(1)}}\mathrm{d}zq\left(r^{(1)}-z\right)\left(r^{(1)}-z\right)q(z)\\&+\frac{g^2}{2}\int_{-\infty}^t\mathrm{d}t'\left(\vec{\alpha}\cdot\hat{r}^{(1)}\right)\delta(t-t')\int_0^{r^{(1)}}\mathrm{d}zq\left(t',r^{(1)}-z\right)q(t',z)\end{aligned} \tag{21}$$

$$\begin{aligned}i\frac{\partial}{\partial t}q\left(r^{(2)}\right)=&-i\left[\vec{\alpha}\cdot\vec{\nabla}_{r^{(2)}},q\left(r^{(2)}\right)\right]-\frac{g^2}{2}\int_0^{r^{(2)}}\mathrm{d}zq\left(r^{(2)}-z\right)\left(r^{(2)}-z\right)q(z)\\&+\frac{g^2}{2}\int_{-\infty}^t\mathrm{d}t'\left(\vec{\alpha}\cdot\hat{r}^{(2)}\right)\delta(t-t')\int_0^{r^{(2)}}\mathrm{d}zq\left(t',r^{(2)}-z\right)q(t',z)\end{aligned} \tag{22}$$

Thus, the WF of the $\pi^+-\pi^0\left(\pi^-\right)$ pair, χ_{pair}, is described as

$$\chi_{pair}=c_1\chi_\pi\left(r^{(1)}\right)+c_2\chi_\pi\left(r^{(2)}\right) \tag{23}$$

where

$$\chi_\pi(r) = const\frac{e^{-\frac{g^2 L_1}{8}r^2}}{r} \tag{24}$$

We obtained the pion WF (Equation (24)) in our prior analysis [26].

Therefore, the magnetization density function and the basis of the electric charge density function of a proton are described as

$$\rho_{mag}^{[P]} = (d_1 m)^2 \left[\frac{e^{-\frac{g^2 L_1}{8}\left(r^{(1)}\right)^2}}{r^{(1)}} + \frac{e^{-\frac{g^2 L_1}{8}\left(r^{(2)}\right)^2}}{r^{(2)}}\right]^2 \tag{25}$$

$$\text{Basis } \rho_{ch}^{[P]} = (d_2 e)^2 \left[\frac{e^{-\frac{g^2 L_1}{8}\left(r^{(1)}\right)^2}}{r^{(1)}} + \frac{e^{-\frac{g^2 L_1}{8}\left(r^{(2)}\right)^2}}{r^{(2)}}\right]^2 \tag{26}$$

here, we combine the latter part of Equation (3) so that the correct magnetization and charge of proton are obtained in Equation (25) and Equation (26). This is because we dropped the factor $\frac{1}{2}$ in the derivation.

By definition, $\vec{r}^{(1)}$ and $\vec{r}^{(2)}$ have the same origin. Thus, considering the direction of momentum $\vec{Q}$ to be the z-axis, $\vec{r}^{(1)}$ and $\vec{r}^{(2)}$ can be expressed by polar coordinates as

$$\vec{r}^{(1)} = \left(r^{(1)}, \theta_1, \phi_1\right), \quad \vec{r}^{(2)} = \left(r^{(2)}, \theta_2, \phi_2\right) \tag{27}$$

Denoting the angle between $\vec{r}^{(1)}$ and $\vec{r}^{(2)}$ as θ and $\vec{r}^{(1)}$ to $\vec{r}$, that is, considering $\vec{r}^{(1)}$ as $\vec{r}$, the magnetization density function of a proton is written

$$\rho_m^{[P]} = (d_1 m)^2 \int_0^\pi \sin\theta_1 \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\phi_1 \left[\frac{e^{-\frac{g^2 L_1}{8}r^2}}{r} + \frac{e^{-\frac{g^2 L_1}{8}(r\cos\theta)^2}}{r\cos\theta}\right]^2 \tag{28}$$

where $\cos\theta = \cos\theta_2 \cos\theta_1 + \sin\theta_2 \sin\theta_1 \cos\left(\phi_2 - \phi_1\right)$

Taking integration to eliminate the θ_2 and ϕ_2 dependence, the actual form of ρ_{mag} can be written as

$$\rho_m^{[P]} = (d_1 m)^2 \left[4\pi \frac{e^{-\frac{g^2 L_1}{4}r^2}}{r^2} + \int_0^\pi \mathrm{d}\theta_2 \int_0^{2\pi} \mathrm{d}\phi_2 \int_0^\pi \sin\theta_1 \int_0^{2\pi} \mathrm{d}\phi_1 \left(2\frac{e^{-\frac{g^2 L_1}{8}r^2\left(1+(\cos\theta)^2\right)}}{r^2} + \frac{e^{-\frac{g^2 L_1}{4}(r\cos\theta)^2}}{r^2}\right)\right] \tag{29}$$

To obtain the electric charge density function of a proton $\rho_{ch}^{[P]}$, we need more careful consideration because of the asymptotic freedom as mentioned in sec 2-3. As $\left|\vec{Q}\right|^2$ becomes larger, r is smaller, $\rho_{ch}^{[P]}$ becomes the summation of two independent $\left|\chi_\pi(r)\right|^2$ terms.

To be precise, $\rho_{ch}^{[P]}$ behaves as

1) as $r \to 0$,

$$\rho_{ch}^{[P]} \to \left(d_2 e\right)^2 2\left(4\pi\right)\frac{\mathrm{e}^{-\frac{g^2 L_1}{4}r^2}}{r^2} \tag{30}$$

2) as $r \to \infty$,

$$\rho_{ch}^{[P]} \to \left(d_1 m\right)^2 \left[4\pi\frac{\mathrm{e}^{-\frac{g^2 L_1}{4}r^2}}{r^2} + \int_0^{\pi}\mathrm{d}\theta_2\int_0^{2\pi}\mathrm{d}\phi_2\int_0^{\pi}\sin\theta_1\int_0^{2\pi}\mathrm{d}\phi_1 \left(2\frac{\mathrm{e}^{-\frac{g^2 L_1}{8}r^2\left(1+(\cos\theta)^2\right)}}{r^2} + \frac{\mathrm{e}^{-\frac{g^2 L_1}{4}(r\cos\theta)^2}}{r^2} \right) \right] \tag{31}$$

The angular integrations become

$$\int_0^{\pi}\mathrm{d}\theta_2\int_0^{2\pi}\mathrm{d}\phi_2\int_0^{\pi}\sin\theta_1\int_0^{2\pi}\mathrm{d}\phi_1 \frac{\mathrm{e}^{-\frac{g^2 L_1}{8}r^2\left(1+(\cos\theta)^2\right)}}{r^2\cos\theta} = 4\pi^3\frac{\mathrm{e}^{-\frac{g^2 L_1}{8}r^2}}{r^2} \tag{32}$$

$$\int_0^{\pi}\mathrm{d}\theta_2\int_0^{2\pi}\mathrm{d}\phi_2\int_0^{\pi}\sin\theta_1\int_0^{2\pi}\mathrm{d}\phi_1 \frac{\mathrm{e}^{-\frac{g^2 L_1}{4}(r\cos\theta)^2}}{\left(r\cos\theta\right)^2} = 4\pi^3\frac{\mathrm{e}^{-\frac{g^2 L_1}{4}r^2}}{r^2} \tag{33}$$

where $\cos\theta = \cos\theta_2\cos\theta_1 + \sin\theta_2\sin\theta_1\cos\left(\phi_2 - \phi_1\right)$

Thus, the magnetization density function and the electric charge density function of a proton are represented by

$$\rho_m^{[P]} = \left(d_1 m\right)^2 4\pi\left[\left(1+3\pi^2\right)\frac{\mathrm{e}^{-\frac{g^2 L_1}{4}r^2}}{r^2}\right] \tag{34}$$

$$\begin{aligned}\rho_{ch}^{[P]} = \left(d_2 e\right)^2 4\pi\Bigg[& 2\frac{\mathrm{e}^{-\frac{g^2 L_1}{4}r^2}}{r^2}\exp\left(-\left(\frac{r^2}{m_P}\right)^{n_P}\right) \\ & + \left[\left(1+3\pi^2\right)\frac{\mathrm{e}^{-\frac{g^2 L_1}{4}(r+\beta)^2}}{\left(r+\beta\right)^2}\right]\left(1-\exp\left(-\left(\frac{r^2}{m_P}\right)^{n_P}\right)\right)\Bigg]\end{aligned} \tag{35}$$

In Equation (35), the $\exp\left(-\left(\frac{r^2}{m_P}\right)^{n_P}\right)$ term shows at what radius asymptotic freedom begins and we treat the m_P, n_P and β values as parameters.

For a neutron, we consider that it is constructed of a $\pi^+ - \pi^-$ pair as mentioned in sec. 2-1. Because π^- is an antiparticle of π^+, the WF of π^- can be considered to be the same as that of π^+. Therefore, the basis of the electric charge density function of a neutron is represented as

$$\begin{aligned}\text{Basis } \rho_{ch}^{[N]} &= \left(h_2\right)^2\left[e\frac{\mathrm{e}^{-\frac{g^2 L_1}{8}\left(r^{(1)}\right)^2}}{r^{(1)}} + \left(-e\right)\frac{\mathrm{e}^{-\frac{g^2 L_1}{8}\left(r^{(2)}\right)^2}}{r^{(2)}}\right]^2 \\ &= \left(h_2 e\right)^2\left[\frac{\mathrm{e}^{-\frac{g^2 L_1}{8}\left(r^{(1)}\right)^2}}{r^{(1)}} - \frac{\mathrm{e}^{-\frac{g^2 L_1}{8}\left(r^{(2)}\right)^2}}{r^{(2)}}\right]^2\end{aligned} \tag{36}$$

Using the same consideration for the vectors $\vec{r}^{(1)}$ and $\vec{r}^{(2)}$ as that of a proton, the basis of the electric charge density function becomes

$$\begin{aligned}\text{Basis } \rho_{ch}^{[N]} &= \left(h_2 e\right)^2 \left[4\pi \frac{\mathrm{e}^{-\frac{g^2 L_1}{4} r^2}}{r^2} - \int_0^{\pi} \mathrm{d}\theta_2 \int_0^{2\pi} \mathrm{d}\phi_2 \int_0^{\pi} \sin\theta_1 \int_0^{2\pi} \mathrm{d}\phi_1 \left(2 \frac{\mathrm{e}^{-\frac{g^2 L_1}{8} r^2 \left(1+(\cos\theta)^2\right)}}{r^2 \cos\theta} - \frac{\mathrm{e}^{-\frac{g^2 L_1}{4}(r\cos\theta)^2}}{\left(r\cos\theta\right)^2} \right) \right] \\ &= \left(h_2 e\right)^2 4\pi \left[\left(1-\pi^2\right) \frac{\mathrm{e}^{-\frac{g^2 L_1}{4} r^2}}{r^2} \right]\end{aligned} \tag{37}$$

For the magnetization density function of a neutron $\rho_m^{[N]}$, the form is the same as the basis of the electric charge density function, but positive. Then $\rho_m^{[N]}$ is represented as

$$\rho_m^{[N]} = \left(h_1 m\right)^2 4\pi \left[\left(\pi^2 - 1\right) \frac{\mathrm{e}^{-\frac{g^2 L_1}{4} r^2}}{r^2} \right] \tag{38}$$

For the electric charge density function of neutron, we again consider asymptotic freedom. At large r (small $\left|\vec{Q}\right|^2$), they move with the same origin, but, at small r (large $\left|\vec{Q}\right|^2$), π^+ and π^- move independently.

To be precise,

3) as $r \to 0$,

$$\rho_{ch}^{[N]} \to \left(h_2 e\right)^2 \left(4\pi\right) 2 \frac{\mathrm{e}^{-\frac{g^2 L_1}{4} r^2}}{r^2} \tag{39}$$

4) as $r \to \infty$,

$$\rho_{ch}^{[N]} \to \left(h_2 e\right)^2 4\pi \left[\left(1-\pi^2\right) \frac{\mathrm{e}^{-\frac{g^2 L_1}{4} r^2}}{r^2} \right] \tag{40}$$

Using the same expression resulting from asymptotic freedom for the proton case, $\rho_{ch}^{[N]}$ is represented by

$$\begin{aligned}\rho_{ch}^{[N]} = \left(h_2 e\right)^2 4\pi \Bigg[& 2 \frac{\mathrm{e}^{-\frac{g^2 L_1}{4} r^2}}{r^2} \exp\left(-\left(\frac{r^2}{m_N} \right)^{n_N} \right) \\ & + \left[\left(1-\pi^2\right) \frac{\mathrm{e}^{-\frac{g^2 L_1}{4}(r+\beta)^2}}{\left(r+\beta\right)^2} \right] \left(1 - \exp\left(-\left(\frac{r^2}{m_N} \right)^{n_N} \right) \right) \Bigg]\end{aligned} \tag{41}$$

where m_N, n_N and β are parameters.

2.3.2. Form Factors

To evaluate the Sachs e.m. FFs of protons and neutrons, *i.e.*, $G_E^{[P]}, G_M^{[P]}, G_E^{[N]}$ and $G_M^{[N]}$, we adopt the following relations proposed by Mitra and Kumari [29].

Accordingly, intrinsic FFs $\tilde{\rho}_m(k)$ and $\tilde{\rho}_{ch}(k)$ are related to the Sachs e.m. FFs $G_M\left(\left|\vec{Q}\right|^2\right)$ and $G_E\left(\left|\vec{Q}\right|^2\right)$ as

$$\tilde{\rho}_{ch}(k) = (1+\tau)^2 G_E\left(\left|\vec{Q}\right|^2\right) \tag{42}$$

$$\mu_{(i)}\tilde{\rho}_m(k) = (1+\tau)^2 G_M\left(\left|\vec{Q}\right|^2\right) \tag{43}$$

where $\tau = \dfrac{\left|\vec{Q}\right|^2}{(2M)^2}$, i = P or N (μ_P and μ_N are the magnetic moment of a proton and a neutron, respectively).

and $\tilde{\rho}(k)$ s are the Fourier transform of the electric charge and magnetization density functions of a nucleon.

Under relativistic consideration, the relationship between k^2 and $\left|\vec{Q}\right|^2$ is

$$k^2 = |\vec{q}|^2 \to \frac{\left|\vec{Q}\right|^2}{1+\tau}$$

and for the nonrelativistic case, the relationship between $k = |\vec{q}|$ and $\left|\vec{Q}\right|$ is $k = |\vec{q}| \to \left|\vec{Q}\right|$.

We derived the electric charge and magnetization density functions in sec 2-3 (a) so that in principle, we only need to take the Fourier Transforms to obtain the Sachs e.m. FFs.

For the magnetization density functions, we can use the Fourier transform directly. However, for the electric charge density functions, we cannot use the exact transformations because that the rigorous Fourier transform cannot reflect the asymptotic freedom characteristics in momentum space. Thus, in the electric charge density function case, we take the Fourier transform of the basis of the electric charge density functions and express the asymptotic freedom in momentum space by adopting a description similar to that used in the configuration space. We then use the relation of Equation (42) to obtain the Sachs FFs of G_E. The electric charge density functions of protons and neutrons were given in sec. 2-3 (a) as

$$\text{Basis } \rho_{ch}^{[P]} = (d_2 e)^2 \left[\frac{e^{-\frac{g^2 L_1}{8}\left(r^{(1)}\right)^2}}{r^{(1)}} + \frac{e^{-\frac{g^2 L_1}{8}\left(r^{(2)}\right)^2}}{r^{(2)}} \right]^2 \tag{44}$$

$$\text{Basis } \rho_{ch}^{[N]} = (h_2 e)^2 \left[\frac{e^{-\frac{g^2 L_1}{8}\left(r^{(1)}\right)^2}}{r^{(1)}} - \frac{e^{-\frac{g^2 L_1}{8}\left(r^{(2)}\right)^2}}{r^{(2)}} \right]^2 \tag{45}$$

Note that, other than the proportional constants, the only difference between Basis $\rho_{ch}^{[P]}$ and Basis $\rho_{ch}^{[N]}$ is the sign.

Considering again the direction of longitudinal momentum $\vec{Q}$ to be the z

axis and considering the polar coordinates $\vec{r}^{(1)}=\left(r^{(1)},\theta_1,\phi_1\right)$ and $\vec{r}^{(2)}=\left(r^{(2)},\theta_2,\phi_2\right)$, and again considering $\vec{r}^{(1)}$ as $\vec{r}$, the Fourier transform of Basis $\rho_{ch}^{[P]}$ and Basis $\rho_{ch}^{[N]}$ can be expressed as

$$\begin{aligned}&\text{Basis }\tilde{\rho}_{ch}^{[P,N]}\left(|\vec{q}|=k\right)\\&=const\,2\pi\int_0^\infty r^2\mathrm{d}r\int_0^\pi\sin\theta_1\mathrm{d}\theta_1\mathrm{e}^{-i|\vec{q}|r\cos\theta_1}\\&\times\left[\mp\frac{\mathrm{e}^{-\frac{g^2L_1}{4}r^2}}{r^2}+\int_0^\pi\mathrm{d}\theta_2\int_0^{2\pi}\mathrm{d}\phi\left(2\frac{\mathrm{e}^{-\frac{g^2L_1}{8}r^2\left(1+(\cos\theta)^2\right)}}{r^2\cos\theta}\mp\frac{\mathrm{e}^{-\frac{g^2L_1}{4}(r\cos\theta)^2}}{\left(r\cos\theta\right)^2}\right)\right]\end{aligned}\tag{46}$$

where $\cos\theta=\cos\theta_2\cos\theta_1+\sin\theta_2\sin\theta_1\cos\left(\phi_2-\phi_1\right)$.

For the second line, we take $\phi_2-\phi_1=\phi$.

The first term of Equation (46) becomes

$$\begin{aligned}\text{First term}&=const\,2\pi\int_0^\infty r^2\mathrm{d}r\int_0^\pi\sin\theta_1\mathrm{d}\theta_1\mathrm{e}^{-i|\vec{q}|r\cos\theta_1}\frac{\mathrm{e}^{-\frac{g^2L_1}{4}r^2}}{r^2}\\&=const\,4\pi\frac{\sqrt{\pi}}{2\sqrt{g^2L_1}}\exp\left(-\frac{|q|^2}{g^2L_1}\right)F\left(1;\frac{3}{2};\frac{|\vec{q}|^2}{g^2L_1}\right)\\&=const\,2\pi\frac{\sqrt{\pi}}{\sqrt{g^2L_1}}F_\pi\end{aligned}\tag{47}$$

where $F\left(\alpha;\beta;z\right)$ is the Kummer's confluent hypergeometric series.

We showed this integral result previously [26].

The second and third terms of Equation (46) become

Fourier cosine of the second term

$$\begin{aligned}&=2const\int_0^\pi\mathrm{d}\theta_2\int_0^{2\pi}\mathrm{d}\phi 2\pi\int_0^\infty r^2\mathrm{d}r\int_0^\pi\sin\theta_1\mathrm{d}\theta_1\cos\left(|\vec{q}|r\cos\theta_1\right)\frac{\mathrm{e}^{-\frac{g^2L_1}{8}r^2\left(1+(\cos\theta)^2\right)}}{r^2\cos\theta}\\&=2const\,\pi\frac{2\pi^2\sqrt{\pi}}{\sqrt{g^2L_1}}\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)\end{aligned}\tag{48}$$

Fourier sine of second term

$$\begin{aligned}&=2const\int_0^\pi\mathrm{d}\theta_2\int_0^{2\pi}\mathrm{d}\phi 2\pi\int_0^\infty r^2\mathrm{d}r\int_0^\pi\sin\theta_1\mathrm{d}\theta_1\sin\left(|\vec{q}|r\cos\theta_1\right)\frac{\mathrm{e}^{-\frac{g^2L_1}{8}r^2\left(1+(\cos\theta)^2\right)}}{r^2\cos\theta}\\&=2cons\,\pi\frac{4\pi^2}{g^2L_1}\sqrt{|\vec{q}|^2}\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)F\left(\frac{1}{2};\frac{3}{2};\frac{|\vec{q}|^2}{g^2L_1}\right)\end{aligned}\tag{49}$$

Fourier cosine of the third term

$$\begin{aligned}&=const\int_0^\pi\mathrm{d}\theta_2\int_0^{2\pi}\mathrm{d}\phi 2\pi\int_0^\infty r^2\mathrm{d}r\int_0^\pi\sin\theta_1\mathrm{d}\theta_1\cos\left(|\vec{q}|r\cos\theta_1\right)\frac{\mathrm{e}^{-\frac{g^2L_1}{4}r^2(\cos\theta)^2}}{\left(r\cos\theta\right)^2}\\&=2const\,\pi\frac{2\pi^2\sqrt{\pi}}{\sqrt{g^2L_1}}\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)\end{aligned}\tag{50}$$

$$\text{Fourier sine og third term} = 0 \tag{51}$$

Therefore, the Fourier transform of the basis of the electric charge density function of protons and neutrons becomes

$$\begin{aligned}&\text{Basis } \tilde{\rho}_{ch}^{[P]}\left(|\vec{q}|\right)\\&= const\,2\pi\frac{\sqrt{\pi}}{\sqrt{g^2L_1}}\left[F_\pi + 4\frac{\pi^{\frac{3}{2}}}{\sqrt{g^2L_1}}\sqrt{|\vec{q}|^2}\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)F\left(\frac{1}{2};\frac{3}{2};\frac{|\vec{q}|^2}{g^2L_1}\right)\right.\\&+ p\frac{1}{|\vec{Q}|}\left[\log\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}} - \exp\left(-\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}}\right)\sum_{n=1}^{\infty}\frac{1}{n!}\left(\sum_{r=1}^{n}\frac{1}{r}\right)\left(\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}}\right)^n\right]\\&\left.+ 4\pi^2\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)\right]\end{aligned} \tag{52}$$

$$\begin{aligned}&\text{Basis } \tilde{\rho}_{ch}^{[N]}\left(|\vec{q}|\right)\\&= const\,2\pi\frac{\sqrt{\pi}}{\sqrt{g^2L_1}}\left|-F_\pi + 4\frac{\pi^{\frac{3}{2}}}{\sqrt{g^2L_1}}\sqrt{|\vec{q}|^2}\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)F\left(\frac{1}{2};\frac{3}{2};\frac{|\vec{q}|^2}{g^2L_1}\right)\right.\\&+ p\frac{1}{|\vec{q}|}\left[\log\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}} - \exp\left(-\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}}\right)\sum_{n=1}^{\infty}\frac{1}{n!}\left(\sum_{r=1}^{n}\frac{1}{r}\right)\left(\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}}\right)^n\right]\\&\left.+ 4\pi^2\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)\right|\end{aligned} \tag{53}$$

Thus,

$$\begin{aligned}&\text{Basis } \tilde{\rho}_{ch}^{[P,N]}\left(|\vec{q}|\right)\\&= const\,2\pi\frac{\sqrt{\pi}}{\sqrt{g^2L_1}}\left[\mp F_\pi + 4\frac{\pi^{\frac{3}{2}}}{\sqrt{g^2L_1}}\sqrt{|\vec{q}|^2}\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)F\left(\frac{1}{2};\frac{3}{2};\frac{|\vec{q}|^2}{g^2L_1}\right)\right.\\&+ p\frac{1}{|\vec{q}|}\left[\log\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}} - \exp\left(-\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}}\right)\sum_{n=1}^{\infty}\frac{1}{n!}\left(\sum_{r=1}^{n}\frac{1}{r}\right)\left(\sqrt{\frac{2|\vec{q}|^2}{g^2L_1}}\right)^n\right]\\&\left.+ 4\pi^2\exp\left(-\frac{|\vec{q}|^2}{g^2L_1}\right)\right]\end{aligned} \tag{54}$$

where p is parameter.

Ye *et al.* used relativistic considerations for their parametrization work [3]. However, in our case, we use nonrelativistic consideration, in which $k = |\vec{q}|$ simply relates to $|\vec{Q}|$. According to Kelly [22], if one knew how to obtain an intrinsic form factor $\tilde{\rho}(k)$ from data for appropriate Sachs form factor, the intrinsic density could be obtained simply by inverting Fourier transform and the naïve nonrelativistic inversion method assumes that $k \to Q$ and $\tilde{\rho}(k) \to G(Q^2)$ where $G(Q^2)$ is the appropriate Sachs form factor. However,

the reason why the nonrelativistic inversion was abandoned is that it produces unsatisfactory results. The corresponding radial densities have an unphysical cusp at the origin and rather hard cores. However, even using relativistic inversin method, this has unphysical failor. Again according to Kelly [22], it is that unique relativistic relationship between the Sachs form factors measured by electron scattering at finite Q^2 and the static charge magnetization densities in nucleon rest frame do not exsit. The basic problem is that electron scattering measures transition matrix elements between state of a composite system that have different momenta and the transition densities between such states are different from the static densities in the rest frame. Firthermore, the boost operator for a composite system depends upon the interaction among its constituents. On the other hand, for our case, we consider that a nuleon is described as a pair of pions which means that the WF has a singularity at the origin. The appearance of a cusp at the origin is rather satisfacotry and is our reason for choosing the nonrelativistic inversion. Thus, from now on, we simply replace $\left|\vec{q}\right|$ with $\left|\vec{Q}\right|$. To construct the electric charge density functions of protons and neutrons in momentum space, we again consider the asymptotic freedom description, just as we did for configuration space. That is, for the small $\left|\vec{Q}\right|^2$ case, the pair of pions moves with respect to the same origin. However, as $\left|\vec{Q}\right|^2$ becomes larger, the two pions begin to move separately and finally move totally independent of each other.

To be precise, this situation is described as

$$\begin{aligned}&\tilde{\rho}_{ch}^{[P]}\left(\left|\vec{Q}\right|\right)\\&= const\, 2\pi\frac{\sqrt{\pi}}{\sqrt{g^2L_1}}F_\pi\left(1+\left(\left(z_1+z_2+z_3\right)/F_\pi\right)\middle/\left(\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_P}\right)^{n_P}\right)\right.\right.\\&\left.\left.+\left(1-\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_P}\right)^{n_P}\right)\right)\frac{z_1+z_2+z_3}{F_\pi}\right)\right)\end{aligned}\tag{55}$$

$$\begin{aligned}&\tilde{\rho}_{ch}^{[N]}\left(\left|\vec{Q}\right|\right)\\&= const\, 2\pi\frac{\pi^{\frac{3}{2}}}{\sqrt{g^2L_1}}F_\pi\left(1-z_3/F_\pi+\left(\frac{\left(z_1+z_2+z_3\right)}{F_\pi}-1\right)\middle/\left(\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_N}\right)^{n_N}\right)\right.\right.\\&\left.\left.+\left(1-\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_N}\right)^{n_N}\right)\right)\left(\frac{\left(z_1+z_2+z_3\right)}{F_\pi}-1\right)\right)\right)\end{aligned}\tag{56}$$

where

$$z_1=4\frac{\pi^{\frac{3}{2}}}{\sqrt{g^2L_1}}\sqrt{\left|\vec{Q}\right|^2}\exp\left(-\frac{\left|\vec{Q}\right|^2}{g^2L_1}\right)F\left(\frac{1}{2};\frac{3}{2};\frac{\left|\vec{Q}\right|^2}{g^2L_1}\right)\tag{57}$$

$$z_2 = p\frac{1}{\sqrt{\left|\vec{Q}\right|^2}}\left[\log\sqrt{\frac{2\left|\vec{Q}\right|^2}{g^2L_1}} - \exp\left(-\sqrt{\frac{2\left|\vec{Q}\right|^2}{g^2L_1}}\right)\sum_{n=1}^{\infty}\frac{1}{n!}\left(\sum_{r=1}^{n}\frac{1}{r}\right)\left(\sqrt{\frac{2\left|\vec{Q}\right|^2}{g^2L_1}}\right)^n\right] \tag{58}$$

$$z_3 = 4\pi^2\exp\left(-\frac{\left|\vec{Q}\right|^2}{g^2L_1}\right) \tag{59}$$

where p of z_2 is parameter.

These expressions are not exactly Fourier transforms of the electric charge density functions. However, because the basis of $\tilde{\rho}_{ch}(k)$ is exactly the Fourier transform of the basis of $\rho_{ch}(r)$, we use the relationship between Sachs e.m. FFs and intrinsic FFs shown in Equation (42) and Equation (43) to obtain in the Sachs e.m. FFs as

$$G_E^{[P]} = \frac{1}{(1+\tau)^2}F_\pi\left(1+\left(\left(z_1+z_2+z_3\right)/F_\pi\right)\middle/\left(\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_P}\right)^{n_P}\right)\right.\right. \\ \left.\left.+\left(1-\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_P}\right)^{n_P}\right)\right)\frac{z_1+z_2+z_3}{F_\pi}\right)\right) \tag{60}$$

$$G_E^{[N]} = \frac{1}{(1+\tau)^2}F_\pi\left(1-z_3/F_\pi+\left(\frac{\left(z_1+z_2+z_3\right)}{F_\pi}-1\right)\middle/\left(\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_N}\right)^{n_N}\right)\right.\right. \\ \left.\left.+\left(1-\exp\left(-\left(\frac{\left|\vec{Q}\right|^2}{m_N}\right)^{n_N}\right)\right)\left(\frac{\left(z_1+z_2+z_3\right)}{F_\pi}-1\right)\right)\right) \tag{61}$$

where z_1, z_2, z_3 are given in Equation (57) to Equation (59), and τ is given as

$$\tau = \frac{\left|\vec{Q}\right|^2}{(2M)^2} \tag{62}$$

here, M is the characteristic mass and it is taken as a parameter.

The relationships between the magnetization density functions and the Sachs e.m. FFs, *i.e.*, $G_M^{[P]}, G_M^{[N]}$, are exactly formulated by their Fourier transform using Equation (43).

Then, we obtain

$$G_M^{[P]} = \frac{1}{(1+\tau)^2}\left(F_\pi\left[1+\left(z_1+z_2+z_3\right)/F_\pi\right]\right) \tag{63}$$

$$G_M^{[N]} = \frac{1}{(1+\tau)^2}F_\pi\left[-1+\left(z_1+z_2+z_3\right)/F_\pi\right] \tag{64}$$

where τ is the same as Equation (59) and z_1, z_2 and z_3 are given in Equa-

tion (57), Equation (58), and Equation (59).

Note that our Sachs e.m. FFs have normalization uncertainty. To compare our values with the parametrization results of Ye *et al.* [3], we have to normalize our $G_M^{[P,N]}$ and $G_E^{[P,N]}$ by dividing them by some constant values. These normalization constants are chosen to be comparable to the above parametrization results. Then the normalized $G_M^{[P]}$ and $G_M^{[N]}$ correspond to the normal $G_M^{[P]}/\mu_P$ and $G_M^{[N]}/\mu_N$, respectively. In section 3 and **Figures 1-8** we use $G_M^{[P]}$ and $G_M^{[N]}$ to denote $G_M^{[P]}/\mu_P$ and $G_M^{[N]}/\mu_N$, respectively.

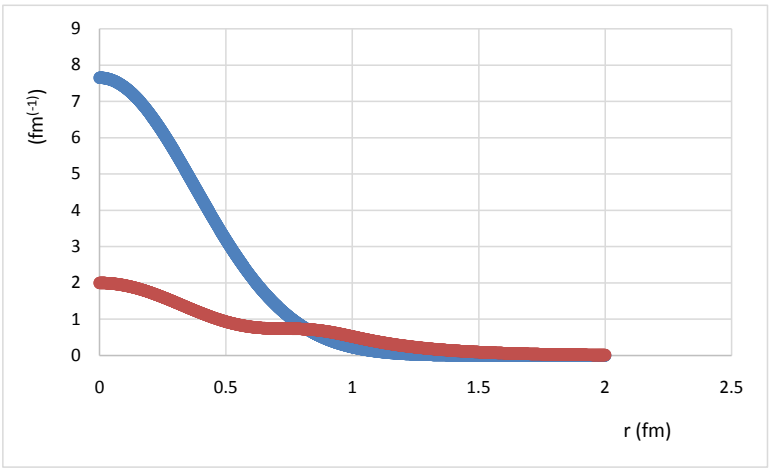


Figure 1. Proton magnetization and electric charge density functions. 1) Blue line is proton magnetization density function (multiplied by r^2) $r^2\rho_{mag}^{[P]}$ (magnitude is $\frac{1}{2}$); 2) Orange line is proton electric charge density function (multiplied by r^2) $r^2\rho_{ch}^{[P]}$. $n_p = 4$; $m_p = 0.7$; $\beta = 0.1$.

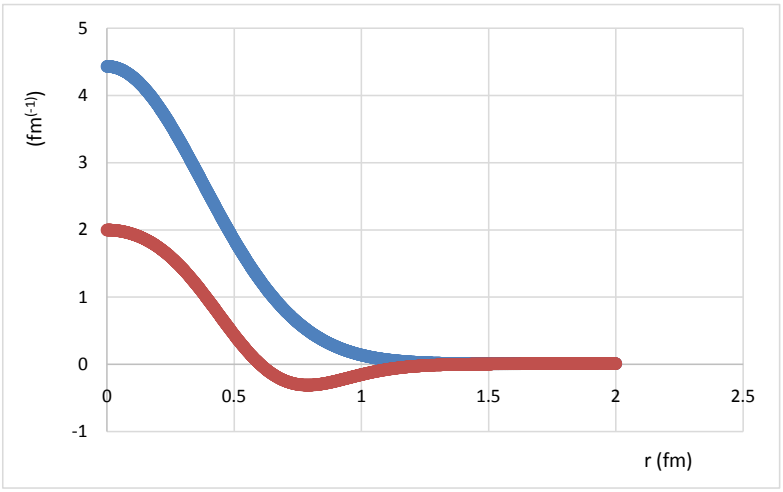


Figure 2. Neutron magnetization and electric charge density function. 1) Blue line is neutron magnetization density function (multiplied by r^2) $r^2\rho_{mag}^{[N]}$; 2) Orange line is neutron electric charge density function (multiplied by r^2) $r^2\rho_{ch}^{[N]}$. $n_N = 4$; $m_N = 0.8$; $\beta = 0.05$.

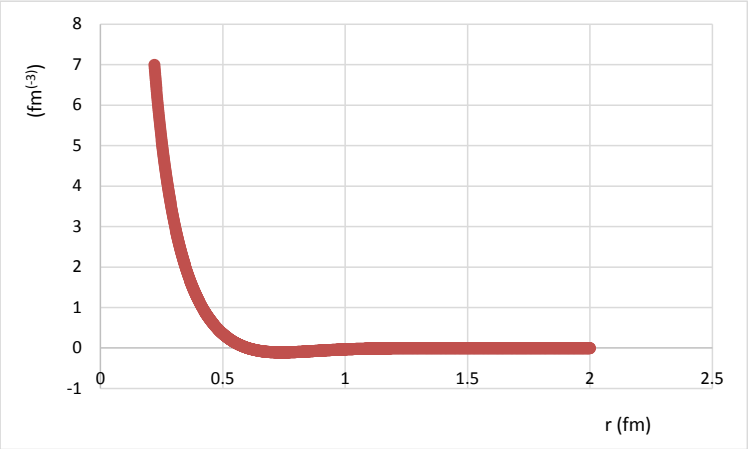


Figure 3. Electric charge density function of neutron $\rho_{ch}^{[N]}$. Note that this becomes negative values (refer **Figure 2**) beyond 0.6 fm region as same as those of Galster model and Maints data beyond 0.7 fm region and that the factor of 1/5 is multiplied.

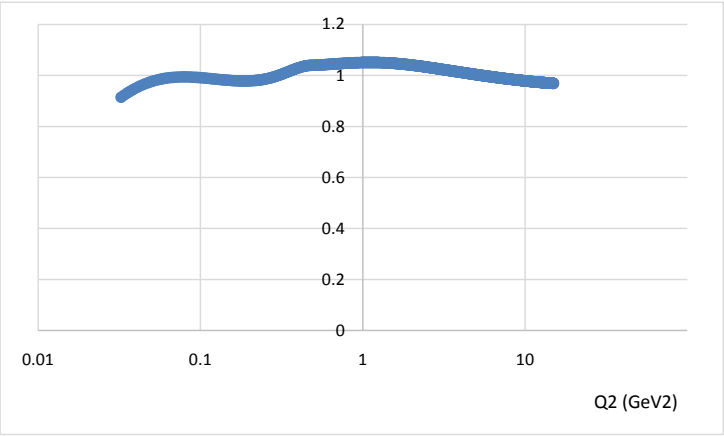


Figure 4. Sachs proton magnetization form factor (divided by G_D) $G_M^{[P]}/G_D$. Parameter p: $p = 3.6284$.

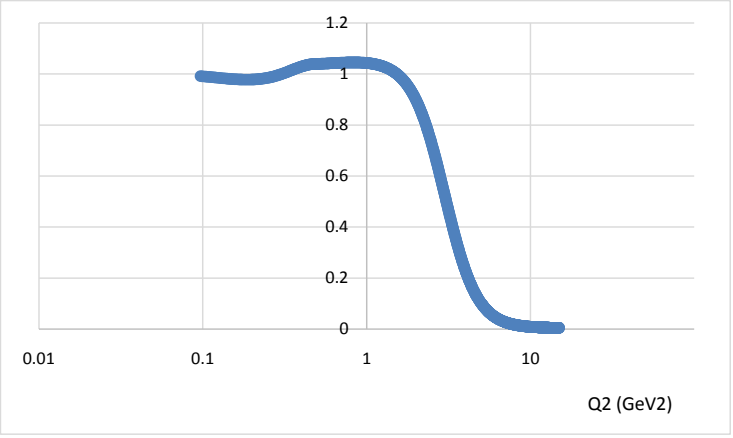


Figure 5. Sachs proton electric charge form factor (divided by G_D) $G_E^{[P]}/G_D$. $n_E^{[P]} = 4$; $m_E^{[P]} = 12$; $p = 3.6284$.

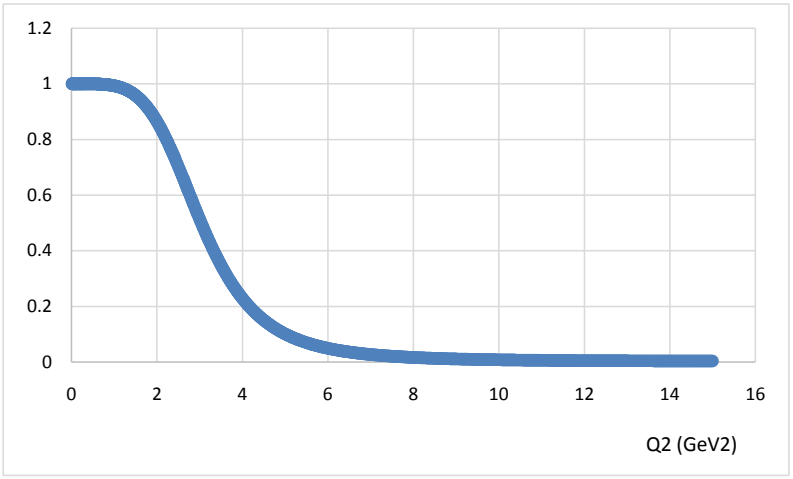


Figure 6. Ratio of $G_E^{[P]}/G_M^{[P]}$.

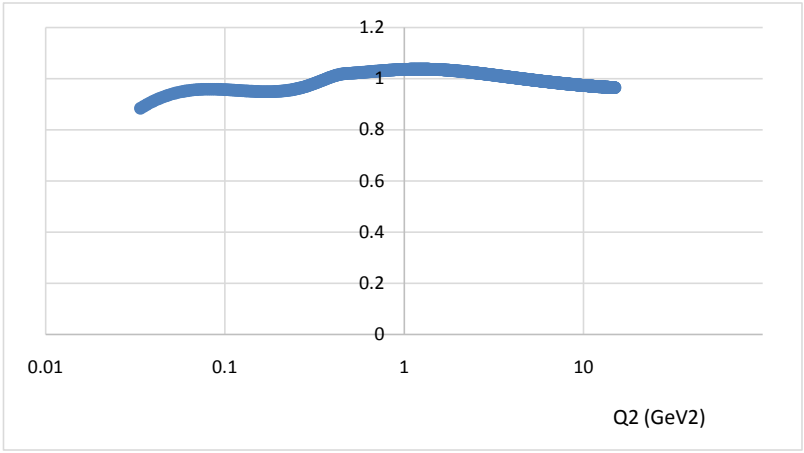


Figure 7. Sachs neutron magnetization form factor (divided by G_D) $G_M^{[N]}/G_D$. Parameter p: $p = 3.6284$.

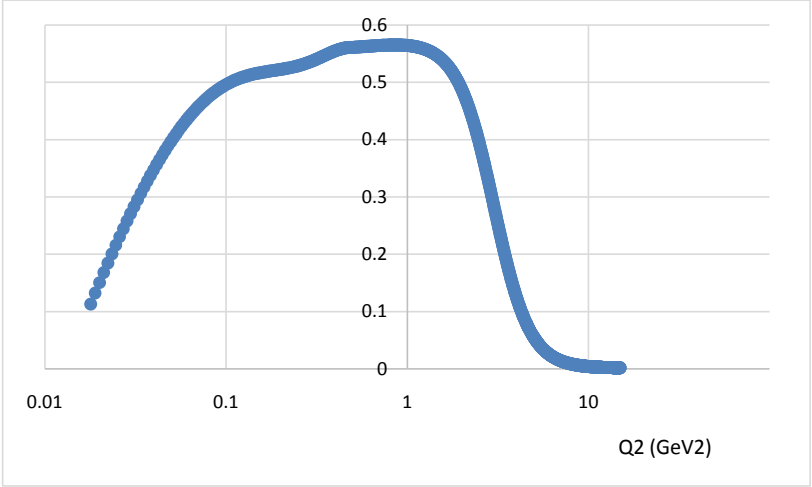


Figure 8. Sachs neutron electric charge form factor (divided by G_D) $G_E^{[N]}/G_D$. $n_E^{[N]} = 4$; $m_E^{[N]} = 12$; $p = 3.6284$.

3. Results

Using Equation (34), Equation (35), Equation (38) and Equation (41), we show the magnetization and electric charge density functions of protons and neutrons in **Figure 1** and **Figure 2**, respectively. Note that the shown density functions are $r^2\rho_m^{[P]}, r^2\rho_{ch}^{[P]}, r^2\rho_m^{[N]}$ and $r^2\rho_{ch}^{[N]}$ instead of just density functions because of the existence of the singularity at the origin (at $r=0$) that results from our definition of the density functions. In addition, we show the electric charge density function of neutron $\rho_{ch}^{[N]}$ in **Figure 3**.

To confine the sizes of protons and neutrons less than 1.2 fm, we chose the Gaussian parameter to be 3.5 (GeV^2). Using this value, we obtain the characteristic mass $\sqrt{\frac{g^2 L_1}{2}}$ of 1025 (MeV), which is similar to the ϕ meson mass. This is different from the pion mass of 140 MeV that we use to evaluate Sachs' proton and neutron e.m. FFs later in this paper.

Figure 1 and **Figure 2** show that our density functions (multiplied by r^2) do not behave exactly like the density functions of Kelly [22], especially, behavior at near origin.

However, Kelly's density functions were obtained by using the relativistic inversion method, which is adopted for preventing them from showing the cusp at origin. To be clear this point, our electric charge density function of neutron in **Figure 3** can be compared to those of Galster model [30] and Mainz data analysed by Schmieden [31]. Their results are shown in Figure 12 of Kelly [22]. Note that their results were obtained using by the nonrelativistic inversion of Fourier transform which is same as our case. Both density functions have singularity at origin although the magnitude of singularity is smaller than ours because their $r^2\rho_{ch}^{[N]}$ goes to 0 as r approach 0. Except the magnitude of singularity, their whole behaviors were quite similar to ours. The density functions of proton also have a similarity such that the proton electric density function overwhelms the proton magnetization density function beyond 0.8 fm. To notify this similarity, we compare our $r^2\rho_{ch}^{[P]}, r^2\rho_m^{[P]}$ with $\rho_{ch}^{[P]}, \rho_m^{[P]}$ in **Figure 5** and **Figure 6** of Kelly's [22] because we are focused in the behavior except in near origin region. This phenomena also appears in the proton electric charge and magnetization density functions of Kelly [22].

Using Equation (60) and Equation (63) with appropriate normalization, we show the results of $G_M^{[P]}/G_D, G_E^{[P]}/G_D$ and $G_E^{[P]}/G_M^{[P]}$ in **Figures 4-6**. Our evaluation forms for Sachs e.m. FFs are not appropriate to show the behavior of the form FFs in the region where $\left|\vec{Q}\right|^2$ is smaller than 10^{-1} (GeV^2). However, they are sufficiently applicable in the region where $\left|\vec{Q}\right|^2$ is larger than 10^{-1} (GeV^2). Thus we can compare our results to the parametrization results in Ye *et al.* [3] in the region where $\left|\vec{Q}\right|^2$ is larger than 10^{-1} (GeV^2). In particular, we obtain a fairly good result for $G_E^{[P]}/G_M^{[P]}$ and also it is quite similar to that of CQM by Miller shown in Arrington [34] up to 4 (GeV^2) as we expected.

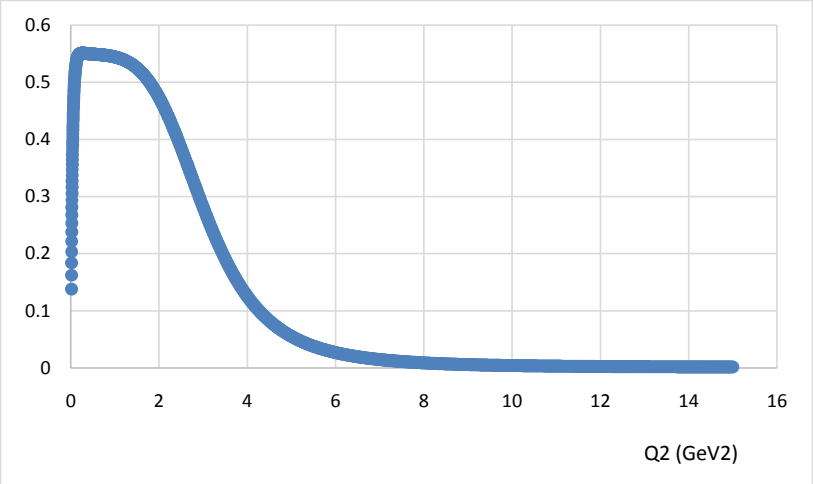


Figure 9. Ratio of $G_E^{[N]}/G_M^{[N]}$.

Using Equation (61) and Equation (64) with appropriate normalization, we show the results of $G_M^{[N]}/G_D, G_E^{[N]}/G_D$ and $G_E^{[N]}/G_M^{[N]}$ in **Figures 7-9**. The magnetization FFs for both protons and neutrons have very similar features to those resulting from parametrization. However, the values of $G_E^{[N]}/G_D$ in the region of 10^{-1} to 10^{0} (GeV^2) are larger than those from the parametrization. Thus, our $G_E^{[N]}/G_M^{[N]}$ shows a faster rising form than it does in other studies [32] [33] [34]. However, the point where our $G_E^{[N]}/G_M^{[N]}$ is the most different is when it converges to zero. This behavior was proposed by Arrington [34], although a reason was not given.

4. Conclusion

We investigate the proton and neutron electromagnetic form factors where the consideration of that nucleon is described as a pion pair. We obtain a good agreement of the electric density function of neutron with Galster model and Maints data except the magnitude of singularity. The density functions of proton also show a similarity to those of Kelly's except near origin. In the case of Sachs e.m. FFs, we obtain a fary good agreement with the parametrization results in Ye *et al.* Therefore, we consider that our description of a nucleon as a pion pair is one of the meaningful aspects.

5. Discussion

As mentioned in conclusion section, we obtain fairy good results in both density functions and Sachs e.m. FFs, however, there is an ambiguous point in our treatment. We do not exactly know the reason why the density functions and the form factors for the magnetization case do not change the form when two pions move independently each other with asymptotic freedom. This may occur because the magnetization arises not as a result of charge distribution, but because of current or spin. Thomas [35] suggests that nucleon spin comes

from the orbital angular momentum of u ($\bar{u}$) and d ($\bar{d}$) quarks (antiquarks). For our case, we describe the nucleon as a pion pair, so we can consider the orbital angular momentum of the u (d) quark and the $\bar{d}$ ($\bar{u}$) antiquark; in the case where the pair of pions has the same origin, angular momentum can arise from movement around the origin. For the independently moving case, the pair of pions move relatively around the center of mass so that the angular momentum can be considered from movement around the center of mass. The important point is that the nucleon spin has a decisive quantity of 1/2 even though the specific orbits of quarks (antiquarks) cannot be determined. Thus, for the case where the magnetization arises because of spin, the magnetization is not affected by the situation of the pair of pions. Conversely, for the case where the magnetization results from current, current arises from the movement of charges, that is, movement of the u (d) quark and the $\bar{d}$ ($\bar{u}$) antiquark, *i.e.*, movement of the pair of pions. Hence we can consider the same argument as the spin case because the magnetization is characterized by the derivative of magnetic energy with respect to the absolute value of the magnetic field at the origin or the center of mass. Either way, we can say that the situation regarding the pair of pions does not affect the magnetization density, but this is not confirmative. To elucidate this point, we need further investigation.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Perdrisat, C.F., Punjabi, V. and Vanderhaeghen, M. (2007) *Progress in Particle and Nuclear Physics*, **59**, 694. https://doi.org/10.1016/j.ppnp.2007.05.001

[2] Arrington, J., de jager, K. and Perdrisat, C.F. (2011) *Journal of Physics: Conference Series*, **299**, Article ID: 012002. https://doi.org/10.1088/1742-6596/299/1/012002

[3] Ye, Z., Arrington, J., Hill, R.J. and Lee, G. (2018) *Physics Letters B*, **777**, 8. https://doi.org/10.1016/j.physletb.2017.11.023

[4] Chung, P.L. and Coester, F. (1991) *Physical Review D*, **44**, 229. https://doi.org/10.1103/PhysRevD.44.229

[5] Schlumpf, F. (1993) *Journal of Physics G*, **20**, 237 (Brief Report) (Relativistic Treatment).

[6] Frank, M.R., Jennings, B.K. and Miller, G.A. (1996) *Physical Review C*, **54**, 920. https://doi.org/10.1103/PhysRevC.54.920

[7] Miller, G.A. and Frank, M.R. (2002) *Physical Review C*, **65**, Article ID: 065205. https://doi.org/10.1103/PhysRevC.65.065205

[8] Cardarelli, F., Pace, E., Salme, G. and Simula, S. (1995) *Physics Letters B*, **357**, 267. https://doi.org/10.1016/0370-2693(95)00921-7

[9] Casptick, S. and Isgur, S. (1986) *Physical Review D*, **34**, 2809. https://doi.org/10.1103/PhysRevD.34.2809

[10] Glozman, L.Y., Papp, Z., Plessas, W., Vaega, K. and Wagenbrunn, R.F. (1998) *Physical Review C*, **57**, 3406. https://doi.org/10.1103/PhysRevC.57.3406

[11] Glozman, L.Y., Plessas, W., Varga, K. and Wagenbrunn, R.F. (1998) *Physical Review D*, **58**, Article ID: 094030. https://doi.org/10.1103/PhysRevD.58.094030

[12] Wagenbrunn, R.F., Boffi, S., Klink, W., Plessas, W. and Radici, M. (2001) *Physics Letters B*, **511**, 31. https://doi.org/10.1016/S0370-2693(01)00622-0

[13] Boffi, S., Glozman, L.Y., Klink, W., Plessas, W., Radici, M. and Wagenbrunn, R.F. (2002) *The European Physical Journal A*, **14**, 17.

[14] Wagenbrunn, R.F., Melde, T. and Plessas, W. (2005) Poincare-Invariant Spectator-Model Currents for Electroweak Nucleon Form Factors.

[15] Gross, F. and Agbakpe, P. (2006) *Physical Review C*, **73**, Article ID: 015203. https://doi.org/10.1103/PhysRevC.73.015203

[16] Miller, G.A. (2002) *Physical Review C*, **66**, 032201(R).

[17] Faessler, A., Gutsche, T., Lyubovitskij, V.E. and Pumsa-ard, K. (2006) *Physical Review D*, **73**, Article ID: 114021. https://doi.org/10.1103/PhysRevD.73.114021

[18] Muller, D., Robschik, D., Geyer, B., Dittes, F.M. and Horejisi, J. (2006) *Fortschritte der Physik*, **42**, 101. https://doi.org/10.1002/prop.2190420202

[19] Ji, X.D. (1997) *Physical Review D*, **55**, 7114. (More thorough Treatment)

[20] Radyushkin, A.V. (1996) *Physics Letters B*, **380**, 417. https://doi.org/10.1016/0370-2693(96)00528-X

[21] Guidal, M., Polyakov, V., Radyushkin, A.V. and Vanderhaegen, M. (2205) *Physical Review D*, **72**, Article ID: 054013. https://doi.org/10.1103/PhysRevD.72.054013

[22] Kelly, J.J. (2002) *Physical Review C*, **66**, Article ID: 065203. https://doi.org/10.1103/PhysRevC.66.065203

[23] Ashman, J., *et al.* (1988) *Physics Letters B*, **206**, 364.

[24] Suura, H. (1978) *Physical Review D*, **17**, 460. https://doi.org/10.1103/PhysRevD.17.469

[25] Suura, H. (1979) *Physical Review D*, **20**, 1429. https://doi.org/10.1103/PhysRevD.20.1412

[26] Kurai, T. (2018) *Results in Physics*, **10**, 865. https://doi.org/10.1016/j.rinp.2018.07.034

[27] Kurai, T. (2014) *Progress of Theoretical and Experimental Physics*, **2014**, 053B01.

[28] Kurai, T. (2017) *Results in Physics*, **7**, 2066. https://doi.org/10.1016/j.rinp.2017.05.028

[29] Mitra, A.N. and Kumari, I. (1977) *Physical Review D*, **15**, 261. https://doi.org/10.1103/PhysRevD.15.261

[30] Galster, S., Klein, H., Moritz, J., Schmidt, D., Wegener, D. and Bleckwenn, J. (1971) *Nuclear Physics B*, **32**, 221. https://doi.org/10.1016/0550-3213(71)90068-X

[31] Schmieden, H. (1999) Form Factors of Neutron. In: Menze, D.W. and Metsch, B., Eds., *Proceeding of the 8th International Conference on the Structure of Baryons*, World Scientific, Singapore, 356.

[32] Bijker, R. and Iachello, F. (2004) *Physical Review C*, **69**, Article ID: 068201. https://doi.org/10.1103/PhysRevC.69.068201

[33] Cloet, I.C., Eichmann, G., El-Bennich, B., Klahn, T. and Roberts, C.D. (2009) *Few-Body*

Systems, **46**, 1. https://doi.org/10.1007/s00601-009-0015-x

[34] Arrington, J. (2016) QCD Bound State Workshop Report Argonne National Lab.

[35] Thomas, A.W. (2008) Interplay of Spin and Orbital Angular Momentum in Proton. arXiv:0803.2775vl [hep-ph]

.

Appendix

Here we show that Gell-mann Nishijima relation still holds under baryon number 0 case.

For mesons, Isospin I, component of Isospin I_3, strangeness S are given as Table 1.

Because Gell-mann Nishijima relation is $Q = I_3 + \frac{1}{2}(B+S)$ (B is baryon number and S is strangeness), this relation holds for meson case because of $B = 0$.

Reminding the fact that field theory shows the duality, we have to add up negative charge of proton p^-, Ξ^+, Ξ^{*+}, Ω^+ to the baryon list. Then using values of Table 1, we can define Table 2 and Table 3 for baryons.

Then it is easy to notice that Gell-mann Nishijima relation also holds for baryon case under the baryon number $B = 0$. This means that baryon number 1 is not necessary.

The verification of the meson pair of each baryons shown in Table 2 and Table 3 are given elsewhere by baryon mass spectra and decay modes.

Table 1. Mesons.

particle	π^+	π^0	π^-	f^0	η^0	k^+	k^-
I	1	1	1	0	0	$\frac{1}{2}$	$\frac{1}{2}$
I_3	1	0	−1	0	0	$\frac{1}{2}$	$-\frac{1}{2}$
strangeness	0	0	0	0	0	1	−1

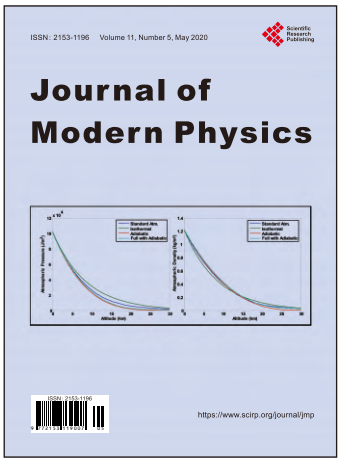
Table 2. Spin 1/2 baryons.

particle	antiparticle	Meson pair	I	I_3	Strangeness
p^+	p^-	$\pi^+ + \pi^0$	1	1	0
$n^0(p^0)$	$n^0(p^0)$	$\pi^+ + \pi^-$	1	0	0
p^-	p^+	$\pi^- + \pi^0$	1	−1	0
Λ^0	Λ^0	$\pi^+ + k^-$ or $\pi^- + k^+$	0	0	0
Σ^+	Σ^-	$\pi^+ + \eta^0$	1	1	0
Σ^0	Σ^0	$\pi^0 + \eta^0$	1	0	0
Σ^-	Σ^+	$\pi^- + \eta^0$	1	−1	0
Ξ^+	Ξ^-	$\pi^0 + k^+$	$\frac{1}{2}$	$\frac{1}{2}$	1
Ξ^-	Ξ^+	$\pi^0 + k^-$	$\frac{1}{2}$	$-\frac{1}{2}$	−1
Ξ^0	Ξ^0	$\pi^+ + k^-$ or $\pi^- + k^+$	0	0	0

Note that we use total sum for Λ^0 and Ξ^0 cases.

Table 3. Spin 3/2 baryons.

particle	antiparticle	Meson pair	I	I_3	strangeness
Δ^+	Δ^-	$\pi^+ + f^0$	1	1	0
Δ^0	Δ^0	$\pi^0 + f^0$	1	0	0
Δ^-	Δ^+	$\pi^- + f^0$	1	−1	0
Σ^{*+}	Σ^{*-}	$\pi^+ + \eta^0$	1	1	0
Σ^{*0}	Σ^{*0}	$\pi^0 + \eta^0$	1	0	0
Σ^{*-}	Σ^{*+}	$\pi^- + \eta^0$	1	−1	0
Ξ^{*+}	Ξ^{*-}	$\eta^0 + k^+$	$\frac{1}{2}$	$\frac{1}{2}$	1
Ξ^{*-}	Ξ^{*+}	$\eta^0 + k^-$	$\frac{1}{2}$	$-\frac{1}{2}$	−1
Ξ^{*0}	Ξ^{*0}	$k^+ + k^-$	0	0	0
Ω^+	Ω^-	$\eta^0 + k^+$	$\frac{1}{2}$	$\frac{1}{2}$	1
Ω^-	Ω^+	$\eta^0 + k^-$	$\frac{1}{2}$	$-\frac{1}{2}$	−1



Call for Papers

Journal of Modern Physics

ISSN: 2153-1196 (Print) ISSN: 2153-120X (Online)
https://www.scirp.org/journal/jmp

Journal of Modern Physics **(JMP)** is an international journal dedicated to the latest advancement of modern physics. The goal of this journal is to provide a platform for scientists and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of modern physics.

Editor-in-Chief

Prof. Yang-Hui He City University, UK

Subject Coverage

Journal of Modern Physics publishes original papers including but not limited to the following fields:

Biophysics and Medical Physics
Complex Systems Physics
Computational Physics
Condensed Matter Physics
Cosmology and Early Universe
Earth and Planetary Sciences
General Relativity
High Energy Astrophysics
High Energy/Accelerator Physics
Instrumentation and Measurement
Interdisciplinary Physics
Materials Sciences and Technology
Mathematical Physics
Mechanical Response of Solids and Structures
New Materials: Micro and Nano-Mechanics and Homogeneization
Non-Equilibrium Thermodynamics and Statistical Mechanics
Nuclear Science and Engineering
Optics
Physics of Nanostructures
Plasma Physics
Quantum Mechanical Developments
Quantum Theory
Relativistic Astrophysics
String Theory
Superconducting Physics
Theoretical High Energy Physics
Thermology

We are also interested in: 1) Short Reports—2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data; 2) Book Reviews—Comments and critiques.

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

https://www.scirp.org/journal/jmp **E-mail: jmp@scirp.org**

What is SCIRP?

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

What is Open Access?

All original research papers published by SCIRP are made freely and permanently accessible online immediately upon publication. To be able to provide open access journals, SCIRP defrays operation costs from authors and subscription charges only for its printed version. Open access publishing allows an immediate, worldwide, barrier-free, open access to the full text of research papers, which is in the best interests of the scientific community.

- High visibility for maximum global exposure with open access publishing model
- Rigorous peer review of research papers
- Prompt faster publication with less cost
- Guaranteed targeted, multidisciplinary audience

Scientific Research Publishing

Website: https://www.scirp.org

Subscription: sub@scirp.org

Advertisement: service@scirp.org