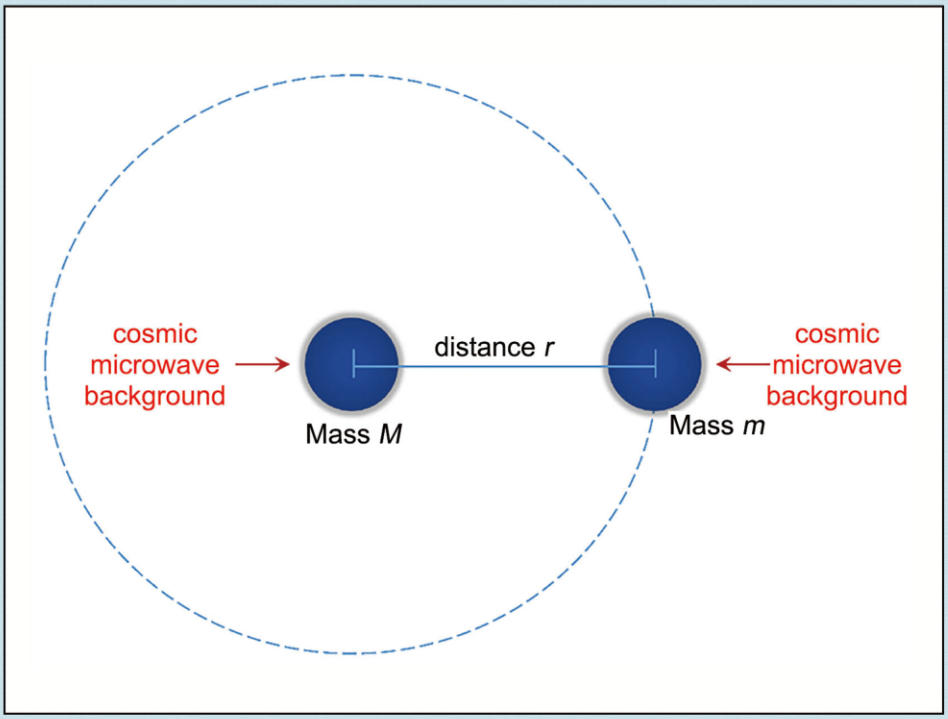
ISSN: 2153-1196 Volume 13, Number 7, July 2022

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, 2022, 13, 1045-1145
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Table of Contents

Volume 13 Number 7 July 2022



The figure on the front cover is from the article published in Journal of Modern Physics, 2022, Vol. 13, No. 7, pp. 1045-1052 by Qinghua Cui.

# 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, 2022, 13, 1045-1052
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Possible Relations of Cosmic Microwave Background with Gravity and Fine-Structure Constant

**Qinghua Cui**

Department of Biomedical Informatics, Peking University, Beijing, China
Email: cuiqinghua@hsc.pku.edu.cn

**How to cite this paper:** Cui, Q.H. (2022) Possible Relations of Cosmic Microwave Background with Gravity and Fine-Structure Constant. *Journal of Modern Physics*, **13**, 1045-1052.
https://doi.org/10.4236/jmp.2022.137058

**Received:** May 29, 2022
**Accepted:** July 3, 2022
**Published:** July 6, 2022



## Abstract

Gravity is the only force that cannot be explained by the Standard Model (SM), the current best theory describing all the known fundamental particles and their forces. Here we reveal that gravitational force can be precisely given by mass of objects and microwave background (CMB) radiation. Moreover, using the same strategy we reveal a relation by which CMB can also precisely define fine-structure constant $\alpha$.

## Keywords

Gravity, Gravitational Constant, Cosmic Microwave Background, Fine-Structure Constant

## 1. Introduction

Gravity, called also gravitational force, represents a fundamental and universal force because of mass and always acts as attraction between all matters. Although dominantly developed by Issac Newton several centuries ago and then by Albert Einstein one century ago, the gravitational force is still one of the most mysterious forces in the universe as it cannot be described by the standard model, the current best theory explaining all the known forces well except gravity [1]. Given the critical role of gravity in the universe and its linear relation with the gravitational constant $G$, one of the most fundamental constants of nature, it is quite important to get the precise knowledge of gravity and $G$. During the past two centuries, more than 300 experimental $G$ values have been measured [2]; however, gravity is still the least precisely known constant due to its extreme weakness and non-shieldability [3].

Here we reveal a quantitative relation between gravity and the cosmic micro-

wave background (CMB), a kind of radiation almost evenly filling the universe [4]. We propose an equation among which the gravitational force and the gravitational constant $G$ can be well determined by the frequency $f$ or temperature $T$ of CMB, which suggests that the gravitational force could be from the interaction between matter and CMB. Based on this finding, the "gravitational waves" are explained as periodic or non-periodic signals of gravity variation in nature. Moreover, this theory can easily interpret why the gravitational force is always attractive and why the "gravitational waves" travel at the speed of light. Moreover, using the same strategy, we reveal that CMB can also precisely predict fine-structure constant $\alpha$.

## 2. Gravitational Force Given by the Cosmic Microwave Background

It is well known that the Boltzmann constant $k$ links the average kinetic energy $E$ of gas particles with its temperature $T$ as the following equation,

$$E = \frac{3}{2}kT \tag{1}$$

It is well known that photons have only two physical degrees of freedom. Thus, the expected average energy of CMB photons will be

$$E = \frac{2}{2}kT = kT \tag{2}$$

Moreover, the energy of CMB photons can be given by

$$E = hf \tag{3}$$

where $h$ is the Planck constant and $f$ is the frequency of the CMB photons. Then, the expected average frequency $f$ of CMB photons is expressed as

$$f = \frac{kT}{h} \tag{4}$$

A widely accepted experimental value of the CMB temperature is $T = 2.73$ K (*Kelvin*) [5], then the expected average frequency of CMB will be $f = 5.688397 \times 10^{10}$.

It is known that gravity is always attractive and gravitational waves travel at speed of light, which triggers us to obtain the idea and hypothesis that CMB could have a role in gravity. As shown in **Figure 1**, we assume there are two astrophysical objects with mass of $M$ and $m$, and with a distance of $r$. It is thus not difficult to understand that CMB crashes into the two objects at all directions except the direction along the line connecting them (**Figure 1**). Therefore, the force by CMB collision will be cancelled out at all directions except the direction along the line connecting them. That is, the inner side of both objects does not receive collision and the corresponding outer side of each object receives the collision (force), thus resulting in an attractive force, that is, the gravitational force, between the two objects along the line connecting them. Then, the gravitational force can be given by the following equation:

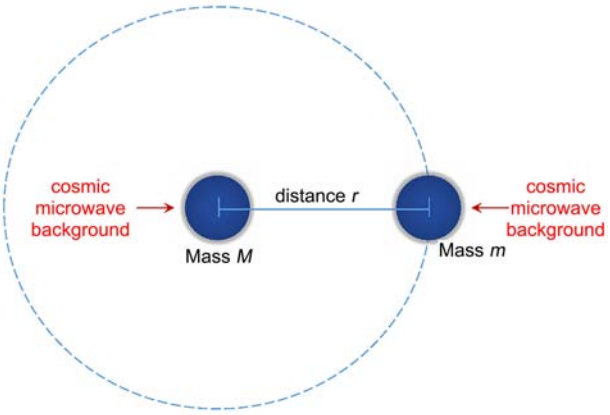



**Figure 1.** A diagram for the role of cosmic microwave background (CMB) in the gravitational force between two astrophysical objects with mass of $M$ and $m$, and with a distance of $r$. In the diagram, CMB plays a role in gravitation as some necessary "stimulus" but not mechanical force of unseen tiny particles as Le Sage's theory of gravitation suggested.

$$F = \frac{Mm}{4\pi r^2} \times \frac{hf}{\lambda} \times f \times 2 \tag{5}$$

where $\frac{Mm}{4\pi r^2}$ is the matter interaction of the two objects. $hf$ and λ are the expected average energy and wavelength of CMB, respectively. Then, $\frac{hf}{\lambda}$ represents the force doing such a work (energy) in a distance of $\lambda$ (collision) by one CMB-wave. And in unit time (second), there are $2f$ times of collisions in unit mass for object pairs. Given that the speed of light in free space is $c = \lambda f$, the above gravitational force equation then can be given by the frequency $f$ or by the CMB temperature $T$ as follows:

$$\begin{cases} F = \dfrac{hf^3}{2\pi c}\dfrac{Mm}{r^2} \\ F = \dfrac{k^3T^3}{2\pi ch^2}\dfrac{Mm}{r^2} \end{cases} \tag{6}$$

Therefore, the Newton's gravitational constant $G$ can be expressed as

$$\begin{cases} G = \dfrac{hf^3}{2\pi c} \\ G = \dfrac{k^3T^3}{2\pi ch^2} \end{cases} \tag{7}$$

As a result, the gravitational constant determined by the current value $T$ and constants $h$, $c$ and $k$ will be $G = 6.474792 \times 10^{-11}$ $m^3 \cdot kg^{-1} \cdot s^{-2}$, which is close to the current measured value ($6.674184 \times 10^{-11}$) [3], however, tension still exists. Given this observation, we have to revisit the temperature of CMB. It should be pointed out that this theory is different from Le Sage's theory of gravitation, which represents the kinetic gravitational theory presented by Nicolas Fatio de Duillier in 1690 and by Georges-Louis Le Sage in 1748. Le Sage's theory ex-

plained Newton's gravitational force as mechanical force of unseen tiny particles, however, in this theory CMB contribute to gravity as some necessary stimulus but not mechanical force as it is too weak to explain gravitation.

On the other hand, using Equation (7) and the current experimental value of $G$, it is not difficult to obtain the corresponding CMB temperature $T = 2.757741$ K. Historically, the experimental and estimated $T$ value does vary [6] [7], for example from [2.64, 2.77] [8] to [2.710, 2.728] [4]. In addition, it is well known that the CMB temperature $T$ displays slight anisotropies [9]. Moreover, it was reported CMB radiation is warmer in the past space-time [10]. For example, $T$ at redshift $z = 2.3371$ could be between 6.0 and 14 K, much higher than the temperature at present space-time [10]. As the cooling of our universe, $G$ continuously decreases, which can easily interpret why astrophysical objects are far away from each other. Moreover, given the non-shieldability of gravity and big distance of astrophysical objects, it makes sense that $T$ at the scale of measured gravity will be a little bit higher than the one (2.73 K) at present space-time. Therefore, for something with shieldability and small measurement scale, it is expected that the temperature of CMB will be a little bit smaller, for example 2.71 K.

In addition, under the proposed framework, we can revisit some famous issue in gravity, e.g. gravitational waves. In Albert Einstein's theory of general relativity, gravity is not a force but modeled as a curvature in space-time. A famous prediction of the general relativity is that accelerating astrophysical objects with mass will produce "gravitational waves" [11], that is, fluctuations in the space-time, which were directly detected for the first time by LIGO recently [12]. Here we try to re-interpret "gravitational waves" using the proposed theory. According to Equation (6), any changes in the parameters ($M$, $m$, $r$, and $T$) will change the gravitational force from original $F$ to new $F'$. Then, the variation of the gravitational force $F_v$ will be

$$F_v = F - F' = \frac{k^3}{2\pi ch^2}\left(\frac{MmT^3}{r^2} - \frac{M'm'T'^3}{r'^2}\right) \tag{8}$$

This equation describes the "gravitational waves". If any value of $M$, $m$, $r$, $c$, or $T$ changes dramatically enough, the "gravitational waves" $F_v$ would be observed. It is thus not difficult to infer that the "gravitational waves" will travel at the speed of the CMB radiation, that is, the speed of light. $F_v$ could be periodic or non-periodic, but obviously in most cases it is non-periodic because the parameters in Equation (10) show less periodicity.

## 3. Fine-Structure Constant Given by the Cosmic Microwave Background

Fine-structure constant α, called also Sommerfeld's constant, is a dimensionless physical constant, which characterizes the strength of the electromagnetic interaction between elementary charged particles [13]. Given its central importance to the foundations of physics, this constant has been measured using various

methods [14] [15] [16], but still remains one of the mysterious constants of nature. Here we try to describe fine-structure constant using CMB using the same strategy as above. Fine-structure constant is defined as the ratio of the tangential velocity of the electron in the lowest-energy orbit of the hydrogen atom to the speed of light. Using the same strategy as above, the force $F_e$ between the electric charges of the electron and the proton can be given by the following equation:

$$F_e = \frac{e^2}{4\pi r^2} \times 2f = \frac{e^2 f}{2\pi r^2} \tag{9}$$

where $e$ is the elementary charge ($e = 1.602176634 \times 10^{-19}$C), $r$ is the Bohr radius ($r = 5.2917721067 \times 10^{-11}$ m), and $f$ is the frequency of CMB. Unlike the gravitational force, electric force is the unit interaction ($e^2/4\pi r^2$) of charges in unit time (second). Therefore, $F_e$ is described by Equation (9) as the $2f$ times (in one second) of unit interactions of paired charges. Besides the electric force, there could also exist the gravitational force $F_g$ between the electron and the hydrogen nuclei, which can be given by Equation (10):

$$F_g = \frac{hf^3}{2\pi c} \frac{m_n m_e}{r^2} \tag{10}$$

where $m_n$ and $m_e$ are the rest mass of the electron ($m_e = 9.109 \times 10^{-31}$ kg) and the hydrogen nuclei ($m_n = 1.67 \times 10^{-27}$ kg), respectively. In addition, the central force $F_c$ for the electron with a tangential velocity $v$ at the Bohr orbit can be given by:

$$F_c = \frac{m_e v^2}{r} \tag{11}$$

In a steady state, $F_e + F_g = F_c$. That is,

$$\frac{e^2 f}{2\pi r^2} + \frac{hf^3}{2\pi c} \frac{m_n m_e}{r^2} = \frac{m_e v^2}{r} \tag{12}$$

Using the present frequency $f$ of CMB and other parameters in Equation (12), we can obtain that the electronic force $F_e$ is much greater ($2.36 \times 10^{39}$ times) than the gravitational force $F_g$. In this case, the force between the electron and the hydrogen nuclei can be precisely given by the electronic force $F_e$. Hence, the Equation (12) can be simplified to the following equation in a high precision.

$$\frac{e^2 f}{2\pi r^2} = \frac{m_e v^2}{r} \tag{13}$$

Then, the tangential velocity $v$ of the electron with the Bohr radius can be given by:

$$v = e\sqrt{\frac{f}{2\pi r m_e}} \tag{14}$$

Then, fine-structure constant will be

$$\alpha = \frac{v}{c} = \frac{e}{c}\sqrt{\frac{f}{2\pi r m_e}} \tag{15}$$

Given the shieldability effect analyzed above, the temperature of CMB in an

atom would be a little bit smaller than the observed one (~2.73 K), for example 2.71 K. As a result, using this CMB temperature (~2.73 K), we have calculated fine-structure constant to be $\alpha^{-1}$ = 137.036804055, which is quite close to the experimental values, for example the values by Smiciklas *et al.* (137.03599955) [17], by Morel *et al.* (137.035999206) [14], by Pachucki *et al.* (137.0360011), by Parker *et al.* (137.035999046), and by Aoyama *et al.* (137.0359991491) [18]. This finding suggests that CMB has a critical connection with fine-structure constant. On the other side, if we take the average (137.03599961) of the above five experimental values, the corresponding temperature of CMB will be ~2.710032 K. In addition, given that the Bohr radius is given by

$$r = \frac{4\pi\varepsilon_0\left(\frac{h}{2\pi}\right)^2}{m_e e^2} \tag{16}$$

where $\varepsilon_0$ is vacuum permittivity. Then, Equation (14) can be further described by

$$\alpha = \frac{e^2}{hc}\sqrt{\frac{f}{2\varepsilon_0}} \tag{17}$$

Moreover, it is well known that $\alpha = e^2/(2\varepsilon_0 hc)$. Therefore, it is not difficult to obtain $\varepsilon_0 = 1/(2f)$, suggesting that vacuum permittivity could be the time used for once stimulus on the interacting charges by CMB. Finally, fine-structure constant can be further expressed as the frequency of temperature of CMB with a number of other constants, as follows:

$$\begin{cases} \alpha = \dfrac{e^2 f}{hc} \\ \alpha = \dfrac{e^2 kT}{h^2 c} \end{cases} \tag{18}$$

From the above analysis, we know that based on the temperature of CMB $T$ = 2.757741 K and $T$ = 2.710032 K, the predicted $G$ value and $\alpha$ value match the experimental values with very high precision. However, at present, the widely accepted temperature is 2.72548 K, indicating that the temperature of CMB is variable at different space-time scale or conditions (e.g. shieldability or non-shieldability) [9]. For example, it was reported that the temperature of CMB at a past time (redshift $z$ = 2.3371) is between 6.0 and 14 K. We summarized the relations of CMB with gravitational constant and fine-structure constant (**Figure 2**).

## 4. Conclusion and Discussion

We have revealed and quantified possible relations of the cosmic microwave background (CMB) with gravity and fine-structure constant. These relations can easily interpret a number of observations. For example, the gravitational force is always attractive and the gravitational wave travels at the speed of light. Moreover, as CMB is continuously cooling, the gravitational constant is expected to decrease as the CMB temperature, which can easily explain why astrophysical objects are far away from each other. In addition, we noted that although the

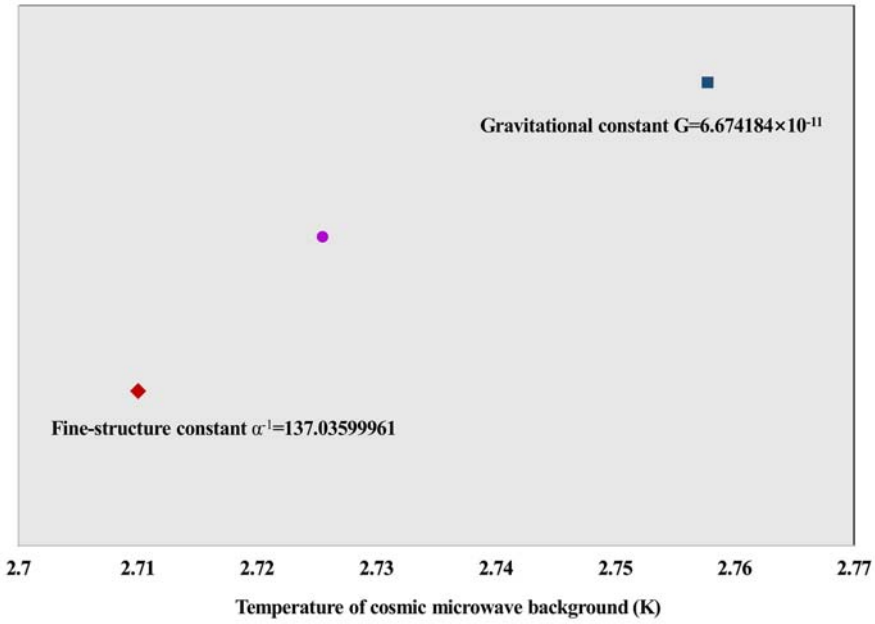



Figure 2. Cosmic microwave background (CMB) can precisely predict the current measured value of fine-structure constant at 2.710032 K and the current measured value of gravitational constant at 2.757741 K. The y-axis value has no meaning but just to separate the data points.

proposed equations match the observation data well but the original dimensions do not match consistently. Therefore, some traditional dimensions should be removed or new dimension should be introduced. Although clear relations of CMB with gravity and fine-structure constant have been revealed, exact explanations are still needed. In addition, it should be noted that besides CMB photons, the cosmic background may also include large amounts of neutrinos and gravitons. Therefore, it is important to investigate whether these "particles" contribute to gravity. Finally, although the proposed equations match the observations well, more data and especially new experiments are needed to further support the proposed findings.

## Acknowledgements

This work has been supported by the grant from the Natural Science Foundation of China (62025102).

## Conflicts of Interest

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

## References

[1] Virdee, T.S. (2016) *Philosophical Transactions of the Royal Society A Mathematical Physical and Engineering Sciences*, **374**, Article ID: 20150259. https://doi.org/10.1098/rsta.2015.0259

[2] Li, Q., *et al.* (2014) *Philosophical Transactions of the Royal Society A Mathematical Physical and Engineering Sciences*, **372**, Article ID: 20140141. https://doi.org/10.1098/rsta.2014.0141

[3] Li, Q., *et al.* (2018) *Nature*, **560**, 582-588. https://doi.org/10.1038/s41586-018-0431-5

[4] Klimenko, V.V., Ivanchik, A.V., Petitjean, P., Noterdaeme, P. and Srianand, R. (2020) *Astronomy Letters*, **46**, 715-725. https://doi.org/10.1134/S1063773720110031

[5] Fixsen, D.J. (2009) *Astrophysical Journal*, **707**, 916-920. https://doi.org/10.1088/0004-637X/707/2/916

[6] Konar, K., Bose, K. and Paul, R.K. (2021) *Scientific Reports*, **11**, Article No. 1008. https://doi.org/10.1038/s41598-020-80195-3

[7] Traunmüller, H. (2020) *F*1000*Research*, **9**, 261. https://doi.org/10.12688/f1000research.22432.3

[8] Amici, G.D., Bensadoun, M., Bersanelli, M., Kogut, A., Levin, S., Smoot, G.F. and Witebsky, C. (1990) *Astrophysical Journal*, **359**, 219. https://doi.org/10.1086/169052

[9] de Bernardis, P., *et al.* (2000) *Nature*, **404**, 955-959. https://doi.org/10.1038/35010035

[10] Srianand, R., Petitjean, P. and Ledoux, C. (2000) *Nature*, **408**, 931-935. https://doi.org/10.1038/35050020

[11] Vitale, S. (2021) *Science*, **372**, eabc7397. https://doi.org/10.1126/science.abc7397

[12] Abbott, B.P., *et al.* (2016) *Physical Review Letters*, **116**, Article ID: 061102.

[13] Müller, H. (2020) *Nature*, **588**, 37-38. https://doi.org/10.1038/d41586-020-03314-0

[14] Morel, L., *et al.* (2020) *Nature*, **588**, 61-65. https://doi.org/10.1038/s41586-020-2964-7

[15] Pachucki, K. and Yerokhin, V.A. (2010) *Physical Review Letters*, **104**, Article ID: 070403. https://doi.org/10.1103/PhysRevLett.104.070403

[16] Parker, R.H., *et al.* (2018) *Science*, **360**, 191-195. https://doi.org/10.1126/science.aap7706

[17] Smiciklas, M. and Shiner, D. (2010) *Physical Review Letters*, **105**, Article ID: 123001. https://doi.org/10.1103/PhysRevLett.105.123001

[18] Aoyama, T., Knoshita, T. and Nio, M. (2018) *Physical Review D*, **97**, Article ID: 036001. https://doi.org/10.1103/PhysRevD.97.036001

Journal of Modern Physics, 2022, 13, 1053-1064
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Gravitational Time Dilation inside the Solid Sphere

**Maciej Rybicki**

Independent Researcher, Kraków, Poland
Email: maciej.rybicki@icloud.com

**How to cite this paper:** Rybicki, M. (2022) Gravitational Time Dilation inside the Solid Sphere. *Journal of Modern Physics*, **13**, 1053-1064.
https://doi.org/10.4236/jmp.2022.137059

**Received:** June 1, 2022
**Accepted:** July 9, 2022
**Published:** July 12, 2022



## Abstract

Gravitational time dilation directly reflects the difference between gravitational potentials at different altitudes in the gravitational field. At the same time this phenomenon is expected to obey the Einstein's equivalence principle, one of two pillars (apart from general covariance) of general relativity. The experiments aimed at detecting the gravitational time dilation are therefore described as the tests of general relativity or, alternatively, the tests of equivalence principle. When applied to the exterior of a solid sphere, these two interpretations are fully compatible both theoretically and experimentally. However, when applied to the interior of a solid sphere (e.g., to the interior of Earth), they seem to contradict each other. Namely, a strict dependence of the gravitational time dilation on the gravitational potential inside the sphere proves to be at odds with the equivalence principle. This paper reveals this problem and provides solution to it. As a consequence, it is concluded that, contrary to the current belief, the Earth's center is older, not younger, than the Earth's surface. Since all the previous experiments have been performed either on or above the Earth's surface, an experiment performed below the Earth's surface is proposed.

## Keywords

Equivalence Principle, General Relativity, G-Force, Gravitational Potential, Gravitational Time Dilation

## 1. Introduction

Gravitational time dilation is a form of time dilation predicted by general relativity (GR), referring to an actual passage of time—the difference of elapsed time between two events as measured at different altitudes in the gravitational field. It relates to the gravitational potentials (metric tensor) at these altitudes. This

phenomenon was originally predicted by Einstein prior to the formulation of GR, as a consequence of applying special relativity to the accelerated frames of reference, hence without direct regard to the gravitational mass (Einstein [1] [2]). It is therefore strictly connected with the equivalence principle, *i.e.*, the Einstein's observation paving the way to GR, according to which there is no experimental difference between the inertial frame of reference and the (local) frame in free fall, as well as between the local reference frame at rest in the uniform in gravitational field (e.g., on the surface of Earth) and the reference frame under uniform acceleration.

Gravitational time dilation has been tested in numerous experiments, to name most important: the Pound-Rebka experiment conducted in 1959 inside the building shaft (tower) of Harvard University (Pound and Rebka [3] [4]); the Hafele-Keating experiment—as a compound effect including both gravitational (due to mass) and kinematic (due to relative velocity) time dilations (Hafele and Keating [5]); Gravity Probe A, performed in 1976 (Vessot *et al.* [6])—the hydrogen maser high precision measurements of the rate of time passage at the altitude of ca. 10,000 km, compared with the measurements of identical maser placed on the Earth's surface; the (Chou *et al.* [7]) experiment with light clocks placed in the Earth's gravitational field, with their altitude differing by only 1 meter. This effect has also a practical relevance: an inclusion of the gravitational time dilation is crucial, apart from the kinematic time dilation, for the correct operation of the GPS (Ashby [8]).

Besides, the experiments aimed at testing the gravitational time dilation in the context of equivalence principle have been conducted using the Mössbauer effect discovered shortly before (Mössbauer [9]). In these experiments, the accelerated system due to rotation of "ultracentrifuge rotor" replaced the gravitational mass (Hay *et al.* [10], Kündig [11]).

Presumably, Richard Feynman was the first who considered this phenomenon in application to the interior of cosmic bodies, specifically to the interior of Earth. According to this great scholar, the inner core (center) of the Earth is, due to the gravitational time dilation, "one or two days" younger than the Earth's crust (surface). Feynman made this illustrative evaluation during his Lectures on Gravitation held at Caltech in 1962/63 (Feynman, Morinigo and Wagner [12]). For a long time taken on trust, the Feynman's estimate has been recently reappraised by Uggerhøj, Mikkelsen and Faye [13], which resulted in its significant correction. Accordingly, the difference between respective ages turned out to be far greater; namely, for the idealized model of the Earth with the assumed uniform density, the center of Earth proved to be 1.58 years younger than the Earth's surface; instead, for the realistic model with factual inhomogeneous mass distribution, this difference increased to 2.49 years. Anyway, no matter if the revealed discrepancy did originate from the Feynman's cursory calculation or from a later misprint in the lecture transcription confusing days with years, the new results differ from the old ones by the magnitude only. Since both estimates share the same theoretical framework, they concordantly state that the Earth's

core is younger than the Earth's crust. There is a general agreement that, due to the gravitational time dilation, the time at the center of Earth (and, generally, of any other massive spherical cosmic body, hereinafter referred to as "solid sphere") is passing slower than on the surface. This stems from the established conviction "probated" by the authority of Feynman, according to which, both outside and inside a solid sphere, gravitational time dilation and gravitational potential are linked by the linear relationship.

A hypothetical clock located at the center of Earth occupies the lowest point of the Earth gravity well, which corresponds with the lowest, negative by convention, gravitational potential. Consequently, the respective clock rate is thought to be the slowest one compared to the rate of any other clock located on the radial path both below and above the surface. This conclusion seems also to follow, as the logical extension, from the experiments aimed at detecting the gravitational time dilation performed on, or above, the Earth's surface. The representative (and earliest) example is the Pound-Rebka experiment. The respective rates of time manifesting themselves through the differences in the gamma ray frequency at different altitudes directly reflect the difference between the gravitational potentials. Although both frequencies were measured outside the sphere (at the top and bottom of the building shaft), the obtained result has been extrapolated by Feynman and his successors (including Uggerhøj) on the whole radial path, both outside and inside the Earth. A direct dependence between gravitational time dilation and gravitational potential in both these cases is treated as obvious. Consequently, the time at the center of Earth is thought to lag behind the surface time, in result of which the Earth's inner core is supposed to be younger than the Earth's crust.

## 2. The Uggerhøj's *et al.* Paper

In introduction to their paper, Uggerhøj, Mikkelsen and Faye (henceforth collectively titled the "Authors") write: "…arguments based on symmetry will convince most skeptics, including those from 'the general public', that there is no gravitational force at the Earth center. Consequently, such an effect [*i.e.*, gravitational time dilation] cannot be due to the force itself, but may instead be due to the 'accumulated action of gravity' (a layman expression for the gravitational potential energy being the radial integral of the force)" [13].

This is a key passage determining the further conclusions of the cited paper. The Authors take for granted that time passes slower at the center of Earth, which unavoidably implies disconnection of the gravitational time dilation from the g-force (interpreted as the proper acceleration). As a consequence, the time dilation must depend directly on the gravitational potential. Hence, in so far as the Authors find justified to reappraise the Feynman's quantitative prediction, they do not intend to question its underlying theoretical framework, determining the general age-relation.

For the sake of transparency (and also with the aim to adopt respective nota-

tion), let us quote almost exactly the formal derivation placed in the first part of the cited paper, concerning the relationship between the gravitational time dilation and gravitational potential inside the Earth, for the homogenous distribution of mass. Accordingly, the gravitational potential ($\Phi$) for the exterior of the solid sphere is

$$\Phi_{ext} = -G\frac{M}{r}, \quad r \geq R \tag{1}$$

*G*—Newton's gravitational constant, *M*—sphere mass, *R*—sphere radius, *r*—distance from the center. Instead, the gravitational potential inside, *i.e.*, in the interior of the sphere is

$$\Phi_{int} = -G\frac{M\left(3R^2 - r^2\right)}{2R^3}, \quad r \leq R \tag{2}$$

These two different expressions share the common result at $r = R$ (at the surface):

$$\Phi\left(R\right) = -\frac{GM}{R} \tag{3}$$

Instead, at the center, *i.e.*, for $r = 0$, one has:

$$\Phi\left(0\right) = -\frac{3GM}{2R} \tag{4}$$

The respective difference is therefore:

$$\Delta\Phi = \Phi\left(R\right) - \Phi\left(0\right) = \frac{1}{2}G\frac{M}{R} \tag{5}$$

Consequently, "a difference in gravitational potential implies a time dilation at the point with lower potential" [13], given by the standard gravitational redshift:

$$\omega = \omega_0\left(1 - \frac{\Delta\Phi}{c^2}\right) \tag{6}$$

where $\omega$ and $\omega_0$ are the angular frequencies at the center and at the surface, respectively. Combining Equation (6) with the result of Equation (5), and considering $\Delta\omega = \omega - \omega_0$, gives the difference in the frequencies related to the difference between the gravitational potentials:

$$\Delta\omega = -\frac{1}{2}G\omega_0\frac{M}{Rc^2} \tag{7}$$

## 3. Gravitational Time Dilation and the Equivalence Principle

Let us precisely consider the application of the equivalence principle to our problem. This principle clearly states that, if any effect (hence also the effect of time dilation) takes place in the non-inertial frame due to gravity, it must also take place in the non-inertial frame due to the kinematically determined acceleration. And *vice versa*. Accordingly, an isolated non-inertial observer located, say, in the "windowless box" is basically (*i.e.*, assuming the box small enough to make the tidal forces negligible) unable to detect if the perceived effect is due to

"real" gravity or due to "pseudo-gravity" caused by the engine running. The specific examples of pseudo-gravity are the non-inertial systems due to rotation about an axis, e.g., the rotating toroidal spaceship (an idea exploited in the sci-fi movies so far), the hypergravity centrifuge—the lab device also used for pilots and astronauts training, or the ultracentrifuge rotor used in the experiments testing the equivalence principle in the context of gravitational time dilation. The respective centrifugal acceleration, perceived as g-force, is particularly evocative because it eliminates the relative motion between any pair of two radially positioned clocks aimed at comparing the time rates. A formal condition to make this possible consists in applying the rotating frame, so to say comoving with the centrifuge. At the same time, in the stationary lab frame, inertial by assumption, the time dilation takes place due to the linear orbital motion of a given point laid on the centrifuge arm (edge of the rotor), hence it is the SR time dilation. The time dilation measured in the rotating frame must be identical to the time dilation measured in the stationary lab frame. This is because in both frames this effect is absolute; in the rotating frame as it were by definition, and in the lab frame on the same basis as it is predicted by the twin paradox and verified in practice in the Hafele-Keating experiment. Likewise, in both frames, the clock located at the centrifuge pivot can be recognized as the reference clock with the null time dilation.

The particular question is whether and how the pseudo-gravity due to rotation of the centrifuge is similar to the real gravity due to the gravitational mass. According to the equivalence principle, this similarity is both exact and limited. Namely, apart from the demand of locality (in the case of centrifuge, the pseudo-tidal forces are even much more distinct), a striking difference is that proper acceleration (g-force) on the surface of a planet is centripetal, whereas the g-force perceived on the rotating arm is centrifugal. Therefore, of course, the location of the mass-center cannot be identified with the location of the centrifuge axis. Let us consider this more specifically.

Let $K$ be the stationary lab frame, and $K'$ the centrifuge rotating frame. Let $O$ be the central point of the frame $K'$, coincident with the pivot of centrifuge. Let $E$ be the point at the outer end of the centrifuge arm. Let $L$ be the distance between $O$ and $E$, obviously equal in both frames. The point $O$ represents both the rotating frame $K'$ (as its unique point) and the stationary frame $K$ (as an exemplary point). According to the SR time dilation applied to the frame $K$, the clock located at $E$ goes slower than the clock located at $O$. According to GR (on the base of equivalence principle), the numerically identical effect takes place in the rotating frame $K'$.

Let $v$ be the linear velocity of $E$ in the frame $K$. The SR time dilation in the $K$ is

$$\gamma_{(K)} = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} \tag{8}$$

The gravitational time dilation in the frame $K'$, as compared to the non-dilated reference clock at the center (pivot) is defined as

$$\gamma_{(K')} = \left(1 - \frac{|\Phi|}{c^2}\right)^{-1/2} \tag{9}$$

The gravitational potential and gravitational acceleration relate to each other as

$$|\Phi| = g_{(r)} r \tag{10}$$

According to the equivalence principle, the centrifugal acceleration at point $E$, being

$$a_c = v^2 / L \tag{11}$$

can be considered equivalent to the centripetal gravitational acceleration (due to the presence of gravitational mass):

$$g_{(r)} = \left| -GM / r^2 \right| \tag{12}$$

Consequently, also the time dilation factors $\gamma_{(K)}$ and $\gamma_{(K')}$ would be equivalent (equal in value). This can only be achieved if we identify $L$ with $r$. Then, by multiplying $a_c \times L$ and $g_{(r)} \times r$, we would identify $v^2$ with $g_{(r)} r$. Can we do that? The short answer is: yes, because time dilations in both $K$ and $K'$ are absolute. And since we deal with the same pair of clocks, both factors have to be identical. However, as far as the purpose to identify $L$ with $r$ is clear, it is not as much clear the reason (possibility) for doing that. For example, the gravitational acceleration (g-force) on the Earth's surface, unitary by convention, is ~9.8 $ms^{-2}$ with the Earth radius being $6.37 \times 10^6$ m, whereas the same acceleration 1g (and much greater) can be easily obtained using the centrifuge with the arm few meters long only, or the ultracentrifuge with the radius few centimeters only. So, it follows that equal accelerations can be associated with extremely different radii. Hence, how $L$ and $r$ can be identified?

The answer is pretty trivial. Although the equivalence principle implies deep consequences leading to general relativity, we don't need to dig into the GR details. The equality between $L$ and $r$ is taken by assumption, whereas the remaining quantities: either mass or linear velocity, should be considered as variables that have to be adjusted to obtain given preset value of acceleration. There are two options, basically. If we start with the centrifuge arm of definite length, then, to equalize the radius connected with gravity with the arm length, the gravitational mass has to be adjusted to match the preset acceleration. If, in turn, we start with the definite mass and radius due to gravity, then, in order to assume the same length of the centrifuge arm, we have to adjust the linear velocity to match the preset acceleration.

In general (*i.e.*, regardless of the details discussed above), the basis for identifying the "pseudo-gravity" due to kinematic acceleration with the "real" gravity due to gravitational mass is the equivalence principle. This means however that the reason for which the clock located at the outer end of the centrifuge rotating arm lags behind the clock located close to the pivot is that it perceives the centrifugal force indistinguishable from the force of gravity. In both cases, there is

one and the same g-force, in formal terms the proper acceleration measurable by accelerometer. The term "perceive" (roughly tantamount to "feel" or "sense") has an unambiguous physical meaning; e.g., exceeding certain critical value of the centrifugal acceleration would result in the damage of clock or, in the case of a trained pilot, in the loss of consciousness. Consequently, if a clock does not actually "perceive" any g-force, one cannot expect it to go slower, hence to undergo the gravitational time dilation. This apparently obvious claim, based on the equivalence principle, has a crucial importance to our problem.

At the center of Earth defined by symmetrical distribution of mass, the lowest gravitational potential coincides with zero g-force. Therefore, to obey the equivalence principle, we shouldn't expect gravitational time dilation to occur there. In fact, it doesn't matter if the clock is located in the center of a planet or in "empty space" far away from any gravity sources. In other words, it is not important whether the g-force is "actually" absent or if it is only "effectively" absent—being neutralized due to the generally conceived free fall (the motion along geodesic), ranging between the rectilinear accelerated motion along the radius and the orbital motion with constant linear velocity, the latter including specific case of a body remaining at rest in any of the five Lagrange points. In both "actual" and "effective" cases, the onboard accelerators (and clocks) do not perceive any g-force, which eventually implies the lack of gravitational time dilation. Otherwise, the equivalence principle would be nothing but a groundless demand. An obvious precondition for the equivalence principle to be valid is the requirement that identical g-forces make two local frames (hence clocks) identical with regard to gravity.

## 4. Gravitational Time Dilation near the Event Horizon of the Schwarzschild Black Hole

The gravitational properties of the black hole observed from a distance do not basically differ from these of other cosmic bodies. The differences become important only near the event horizon and beyond. The ratio between the time rate near the event horizon of a non-rotating uncharged black hole and the time rate indicated by remote clock is given by equation:

$$\frac{\Delta\tau}{\Delta t}=\left(1-\frac{r_S}{r}\right)^{1/2} \tag{13}$$

$\Delta\tau$—elapsed proper time between two events close to observer located near the event horizon of black hole; $\Delta t$—elapsed coordinate time between these same events, measured by distant observer; $r_S$—Schwarzschild radius; $r$—radial distance from the center of black hole, provided $r > r_S$. Considering $r_S = 2GM/c^2$, it follows:

$$\frac{\Delta\tau}{\Delta t}=\left(1-\frac{2GM}{c^2 r}\right)^{1/2} \tag{14}$$

The gravitational time dilation factor is therefore:

$$\gamma = \left(1 - \frac{2GM}{rc^2}\right)^{-1/2} \tag{15}$$

This can be alternatively expressed in terms of escape velocity $v_e = \left(2GM/r\right)^{1/2}$, written as the fraction of $c$, *i.e.*, $\beta_e = v_e/c$:

$$\gamma = \left(1 - \beta_e^2\right)^{-1/2} \tag{16}$$

## 5. Gravitational Time Dilation outside and inside the Solid Sphere

Our goal is to reconcile the two seemingly contradictory premises: 1) According to GR, the gravitational time dilation is modeled by the metric tensor, which means that it directly depends on the gravitational potential. This prediction has been confirmed by all previous experiments; 2) Extending this prediction to the interior of a solid sphere violates the equivalence principle. Namely, it contradicts the demand according to which an exemplary windowless box (lab) located at the center of a solid sphere should not differ physically from the identical windowless box located far away from the sources of gravity. This is because, in both cases (in both local inertial frames), the g-force amounts to zero. The previous solution to this problem, represented both by Feynman and Uggerhøj *et al.*, so to say "ignores" the equivalence principle in the application of the gravitational time dilation to the interior of solid sphere. Below, it is proposed an alternative solution consistent with this principle.

Let $R$ be the radius of solid sphere, $M$—mass of this sphere, $r$—radial distance from the center of a basically free magnitude, either greater or less than $R$. Let us consider first the gravitational potential $\Phi$ as the function of $r$. Due to different (regarding gravity) physical conditions inside and outside the solid sphere, the respective relationship is plotted by two separate functions connected at "inflection point" $r = R$, at which $\Phi_{ext} = \Phi_{int} = -GMR^{-1}$. Hence, the overall dependence of $\Phi$ on $r$ takes the form of a single graph consisting of two functions on two complementary half-open intervals, according to the conditions specified on the right sides of Equations (1) and (2). At $r = 0$, the function based on Equation (2) reaches the minimum (lowest gravitational potential): $\Phi_{int} = -\frac{3}{2}GMR^{-1}$. The gravitational potential defined according to Equation (2) determines the shape of the gravity well, for $0 \le r \le R$ (**Figure 1**).

In turn, the gravitational accelerations inside and outside the sphere of radius $R$ are:

$$g_{(r)ext} = -\frac{GM}{r^2}\hat{\mathbf{r}} = \frac{GM}{r^2} \quad \left(r \ge R\right) \tag{17}$$

$$g_{(r)int} = \frac{GM}{R^3}r \quad \left(r \le R\right) \tag{18}$$

($\hat{\mathbf{r}}$—the unity vector directed outward). As previously, the respective graph consists of two functions on two complementary half-open intervals "glued together" at $r = R$ (**Figure 2**).

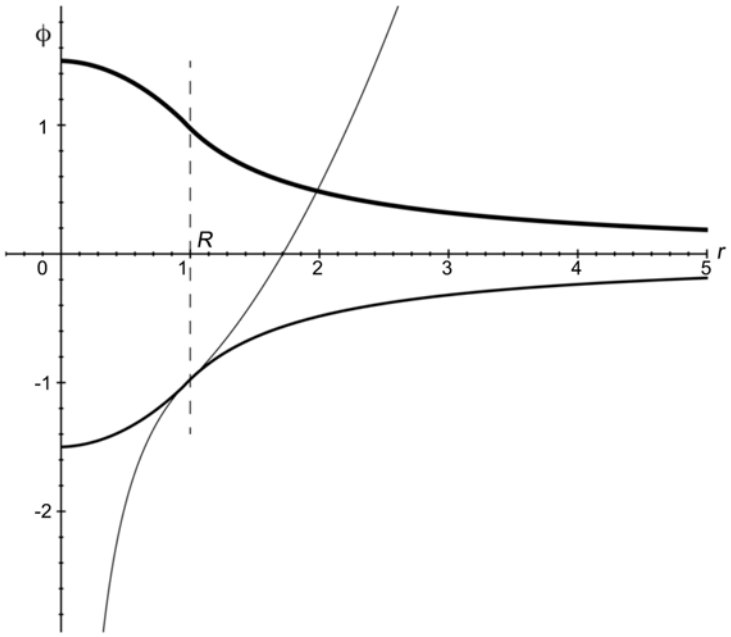



**Figure 1.** Gravitational potential Φ inside and outside the massive spherical body, as a function of distance ($r$). The respective graph (lower quadrant, bold line) consists of two functions on two complementary half-open intervals connected at $r = R$. The upper quadrant graph depicts the gravitational potential absolute value, thought to correspond directly to the gravitational time dilation. The constant factors $G$, $M$ and $R$ are here normalized to unity.

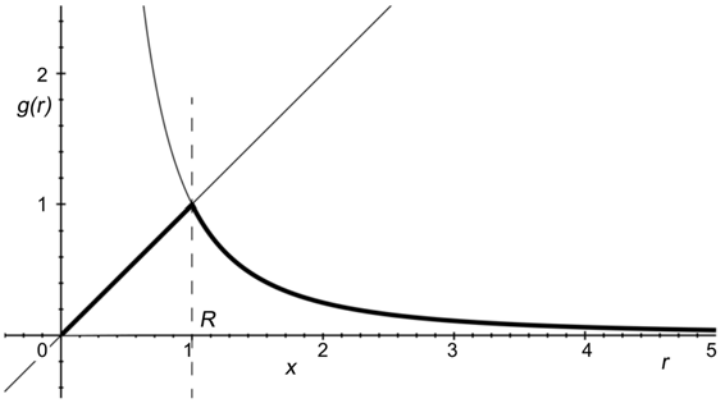



**Figure 2.** Gravitational proper acceleration $g(r)$ (g-force) inside and outside the solid sphere. The respective graph consists of two functions on the two complementary half-open intervals connected at $r = R$. All constant factors are here normalized to unity.

The gravitational time dilation based on the gravitational acceleration, is, in the general case:

$$\gamma = \left(1 - \frac{|\Phi|}{c^2}\right)^{-1/2} \tag{19}$$

This equation is basically consistent with the gravitational time dilation for

the Schwarzschild black hole defined by Equation (13). It is to be noted that gravitational potential Φ is not a quantity directly "perceived". It is rather a mathematical being corresponding to the GR metric tensor. However, the gravitational potential can be factorized into the proper gravitational acceleration (g-force) and the radius. As mentioned before, g-force is a quantity directly "perceivable" by the accelerometer.

Let's rewrite then the gravitational potential as $|\Phi| = g_{(r)}r$ (as we already did in Equation 10). For any case specified as $r \geq R$, this way of defining the gravitational potential is equivalent to the one given by Equation (1), *i.e.*, $\Phi = -G(M/r)$. Hence, outside the sphere, this way of denoting (and defining) Φ is a purely formal operation with no physical consequences. However, for $r \leq R$, things look different. Combining Equation (19) with Equation (10) gives:

$$\gamma = \left(1 - \frac{g_{(r)}r}{c^2}\right)^{-1/2} \tag{20}$$

The next step is the following. We specify $g_{(r)}$ according to Equations (17) and (18) and plug them into the above equation. In result, we obtain the two differing modes to obtain the time dilation factors: one for the exterior and the other one for interior of a solid sphere, both complying with the equivalence principle:

$$\gamma_{ext} = \left(1 - \frac{GM}{c^2 r}\right)^{-1/2} \quad (r \geq R) \tag{21}$$

$$\gamma_{int} = \left(1 - \frac{GMr^2}{c^2 R^3}\right)^{-1/2} \quad (r \leq R) \tag{22}$$

In agreement both with theory and experiments, Equation (21) remains physically identical with Equation (19), expressing linear dependence (proportionality) between the gravitational potential and the gravitational time dilation. Instead, in Equation (22), the factorized gravitational potential makes the actual radius mathematically interacting with the sphere radius, in result of which the gravitational time dilation does not depend any more directly on the gravitational potential. In particular, for $r \to 0$ one has $\gamma(t)_{int} = 1$ (which is also directly obvious considering $g_{(r)} = 0$ at the center). This corresponds with $\gamma(t)_{ext} = 1$ for $r \to \infty$. Hence, the "windowless box" at the center of a solid sphere and the "windowless box" far away in the empty space prove to be identical with each other with regard to the gravitational time dilation, in compliance with the equivalence principle.

## 6. Numerical Estimation for the Idealized Model of Earth

Let us use the equations obtained in the previous section to estimate the difference in age between the Earth's center and Earth's surface, for the simplified model of Earth with the uniform density. At the center of Earth, *i.e.*, at $r = 0$, Equation (22) reduces to $\gamma_{int} = 1$. Instead, at the surface, *i.e.*, for $r = R$, the

Equations (21) and (22) reduce to a one common equation, namely:

$$\gamma_{\text{surface}} = \left(1 - \frac{GM}{c^2 R}\right)^{-1/2} \tag{23}$$

Let us denote: $T_E$ —the overall reference age of Earth, identified with the age of the Earth's center ( $T_{\text{center}}$ ); $T_{\text{surface}}$ —age of the Earth's surface; $\Delta T = T_{\text{center}} - T_{\text{surface}}$ . Hence, the difference is given by:

$$\Delta T = T_E - \frac{1}{\gamma_{\text{surface}}} \times T_E \tag{24}$$

It follows:

$$\Delta T = T_E - \left(1 - \frac{GM}{c^2 R_E}\right)^{1/2} \times T_E \tag{25}$$

Substituting: $c^2 \approx 8.99 \times 10^{16}\ \text{m}^2 \cdot \text{s}^{-2}$; $G \approx 6.67 \times 10^{-11}\ \text{m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$; $T_E \approx 4.55 \times 10^9\ \text{yr}$ ; $M_E \approx 5.97 \times 10^{24}\ \text{kg}$ ; $R_E \approx 6.37 \times 10^6\ \text{m}$ , we obtain after some arithmetic:

$$\Delta T = T_{\text{center}} - T_{\text{surface}} \approx 1.58\ \text{yr} \tag{26}$$

It follows that the center of Earth is approximately 1.58 years older than the Earth's surface—the value equal, but reversely attributed, compared to the figure obtained by Uggerhøj *et al.* The term "approximately" refers here to: 1) assumed homogenous distribution of the Earth mass; 2) assumed constancy of both $G$ and $c$ in cosmic time (precisely, invariability of the factor $G/c^2$). It is to be noted that, in the light of various "non-standard" theories/hypotheses such as VSL or Dirac's LNH, the latter is not obvious by itself.

## 7. Conclusions

A principle-based analysis of the problem of gravitational time dilation in the interior of Earth and other cosmic solid spheres reveals fallacy of the current view represented by Feynman, Uggerhøj and other researchers. The quantitative difference between particular estimates eventually appears less important than the incorrectness of the general assumptions commonly shared. Namely, a consistent application of the Einstein's equivalence principle impels us to revise the so far view as to the relationship between the gravitational time dilation and the gravitational potential. It appears that in the case of interior of a solid sphere, the time dilation depends on g-force rather than on the gravitational potential. As applied to the Earth, this means that relation between the ages of Earth's center and Earth's surface is basically different from the one previously formulated; namely, the inner core of the Earth is not younger, but older than the Earth's crust.

As in most cases in physics, an ultimate criterion to settle a given problem is an experiment. To make it happen, we do not need to reach the Earth's core because all that counts here is the general tendency described by Equation (22). The respective test could be basically similar to the Pound-Rebka experiment

except being performed not above but below the Earth's surface. This should be neither too difficult nor expensive, considering that devices needed for that purpose are typically on standard lab equipment, and that any of numerous inactive mines endowed with vertical shaft could be used for that purpose.

## Conflicts of Interest

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

## References

[1] Einstein, A. (1907) Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen. *Jahrbuch der Radioaktivität und Elektronik*, **4**, 411-462. [English Translation: On the Relativity Principle and the Conclusions Drawn from It. In: Stachel, J., *et al.*, Eds., Beck, A., Trans., *The Collected Papers of Albert Einstein*, The Swiss Years: Writings, 1900-1909, Vol. 2, Princeton University Press, Princeton, 252-311.]

[2] Einstein, A. (1912) Lichtgeschwindigkeit und Statik des Gravitationsfeldes. *Annalen der Physik*, **343**, 355-369. [English Translation: The Speed of Light and the Statics of the Gravitational Field. In: Stachel, J., *et al.*, Eds., Beck, A., Trans., *The Collected Papers of Albert Einstein*, The Swiss Years: Writings, 1912-1914, Vol. 4, Princeton University Press, Princeton, 95-106.] https://doi.org/10.1002/andp.19123430704

[3] Pound, R.V. and Rebka Jr, G.A. (1959) *Physical Review Letters*, **3**, 439-441. https://doi.org/10.1103/PhysRevLett.3.439

[4] Pound, R.V. and Rebka Jr, G.A. (1960) *Physical Review Letters*, **4**, 337. https://doi.org/10.1103/PhysRevLett.4.337

[5] Hafele, J.C. and Keating, R.E. (1972) *Science*, **177**, 4044, 168-170. https://doi.org/10.1126/science.177.4044.168

[6] Vessot, R.F.C., *et al.* (1980) *Physical Review Letters*, **45**, 2081-2084. https://doi.org/10.1103/PhysRevLett.45.2081

[7] Chou, C., Hume, D.B., Rosenband, T. and Wineland, D.J. (2010) *Science*, **329**, 1630-1633. https://doi.org/10.1126/science.1192720

[8] Ashby, N. (2002) *Physics Today*, **55**, 41-47. https://doi.org/10.1063/1.1485583

[9] Mössbauer, R.L. (1958) *Zetschrift für Physik A*, **151**, 124-143. https://doi.org/10.1007/BF01344210

[10] Hay, H.J., Schiffer, J.P., Cranshaw, T.E. and Egelstaff, P.A. (1960) *Physical Review Letters*, **4**, 165-166. https://doi.org/10.1103/PhysRevLett.4.165

[11] Kündig, W. (1963) *Physical Review*, **129**, 2371-2376. https://doi.org/10.1103/PhysRev.129.2371

[12] Feynman, R.P., Morinigo F.B., Wagner, W.G., Preskill, J. and Thorne, K.S. (2003) Feynman Lectures on Gravitation. Hatfield, B., Ed. Taylor & Francis Group. https://doi.org/10.1201/9780429502859

[13] Uggerhøj, U.I., Mikkelsen, R.E. and Faye, J. (2016) *European Journal of Physics*, **37**, Article ID: 35602. https://doi.org/10.1088/0143-0807/37/3/035602

Journal of Modern Physics, 2022, 13, 1065-1079
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Surface Waves in a Relativistic Plasma Stream Propagating in a Duct: Kinetic Theory

**Hee J. Lee[1*], Young Kyung Lim[2]**

[1]Kyunggi-do Namyangju-si Hwado-eup Biryong-ro 321, South Korea
[2]Department of Radiation Oncology, National Cancer Center, Goyang, South Korea
Email: ychjlee@yahoo.com

**How to cite this paper:** Lee, H.J. and Lim, Y.K. (2022) Surface Waves in a Relativistic Plasma Stream Propagating in a Duct: Kinetic Theory. *Journal of Modern Physics*, **13**, 1065-1079.
https://doi.org/10.4236/jmp.2022.137060

**Received:** May 18, 2022
**Accepted:** July 9, 2022
**Published:** July 12, 2022



## Abstract

Dispersion relation of surface waves generated by a relativistic plasma stream in an infinite duct surrounded by vacuum is derived by means of relativistic Vlasov equation. The kinematic boundary condition imposed on the distribution function, the specular reflection conditions on the four sides of duct, can be satisfied by placing infinite number of fictitious surface charge sheets spaced by the duct widths. By placing appropriate fictitious surface charge sheets one can effectively deal with the extended electric field introduced in the Vlasov equation and treat kinetically the surface waves in semi-infinite, slab, and duct plasmas on equal ground. The relativistic duct dispersion relation is compared with the earlier non-relativistic surface wave dispersion relation.

## Keywords

Surface Wave, Relativistic Plasma, Duct, Slab

## 1. Introduction

We investigate surface waves generated by a relativistic plasma beam travelling in a duct interfaced with vacuum by using relativistic Vlasov equation. Surface waves propagate along the interface between two different media while being attenuated in the perpendicular direction. Surface waves are the normal modes that are given rise to in bounded plasmas by satisfying the kinetic and the electrodynamic boundary conditions on the interface between two media. The electrodynamic boundary conditions are the connection formulas matching the fields of the two media, which can be mathematically worked out from the governing equations themselves if the density gradient across the plasma and the other side is very steep. "A sharp interface" is synonymous to a theoretically in-

*Former professor of physics at Hanyang University.

finite density gradient across the boundary. In this case, the connection formula can be obtained by "infinitesimal integration" across the interface, which is the operation performed on a certain relevant equation in the manner $\int_{-\epsilon}^{\epsilon}(\cdots)\mathrm{d}x$, where $\epsilon$ is a positive infinitesimal. If the quantity $(\cdots)$ is a perfect differential, this operation yields a non-vanishing surface term that contributes to the connection formula. Usually the surface term is surface charge or surface current, In this way, the well-known electromagnetic and dynamic boundary conditions on the boundary can be derived [1]. We might say that the surface wave in the plasma and the vacuum side wave are two different manifestations of "the same wave" created in an extreme inhomogeneous plasma. The electromagnetic boundary conditions are the same in both non-relativistic and relativistic plasmas.

Integrating the Maxwell equation

$$\nabla\times\boldsymbol{B}=\frac{4\pi}{c}\boldsymbol{J}+\frac{1}{c}\frac{\partial\boldsymbol{E}}{\partial t}$$

over an infinitesimal segment $(-\epsilon,\epsilon)$ across the interface $x=0$, we obtain

$$\left[B_y\right]=\frac{4\pi}{c}J_z^*$$

where $[\cdots]$ signifies the jump across the interface, $\boldsymbol{J}^*=\int_{-\varepsilon}^{\varepsilon}\boldsymbol{J}\mathrm{d}x$ is the surface current. In a cold plasma, the normal component of the electric field is discontinuous by the amount of surface charge $\sigma$: $\left[E_x\right]=4\pi\sigma$. On the other hand, we have the relation $J_z^*=u\sigma$ due to the charge conservation law, where $u$ is the drift velocity of the plasma in the $z$ direction, and $c$ is the speed of light. Thus we have the characteristic boundary condition in a cold streaming plasma [1] [2]:

$$\left[B_y\right]=\frac{u}{c}\left[E_x\right] \tag{1}$$

or equivalently,

$$\left[D_x\right]=\frac{ck_z}{\omega}\left[B_y\right] \tag{2}$$

where $D_x$ is the normal to the interface component of $\boldsymbol{D}$, the electric displacement, $k_z$ is the $z$ component of the wave vector, and $\omega$ is the wave frequency. The casual use of $\left[D_x\right]=0$ or $\left[B_y\right]=0$ in a drifting plasma leads to erroneous results as discussed in earlier works [1] [3]. The physical origin of the boundary relation in Equation (1) is due to the surface current formed in a cold streaming plasma as is evident in the above derivation. Equation (1) is valid in non-relativistic as well as in relativistic plasmas [2].

In bounded Vlasov plasmas, the kinematic boundary condition that is usually referred to as the specular reflection condition is assumed to be satisfied on a sharp boundary [4], regardless of whether the plasma is described by non-relativistic distribution function $f(\boldsymbol{r},\boldsymbol{v},t)$ or relativistic distribution function $f(\boldsymbol{r},\boldsymbol{p},t)$ (see Equation (9) below). Most of the works on bounded kinetic plasmas were dealt with non-relativistically. Non-relativistic kinetic theory of surface waves in semi-infinite plasmas is well-known [5] [6]. Non-relativistic

kinetic dispersion relation of surface wave in a slab plasma was worked out earlier [7].

In this work, we investigate surface waves of a moving relativistic Vlasov plasma in a duct. We consider an infinite duct formed by intersections of four planes: $x=0,a$ and $y=0,b$, with $-\infty<z<\infty$. We shall discuss the specular reflection boundary condition in terms of $f(\boldsymbol{r},\boldsymbol{v},t)$, rather than in terms of $f(\boldsymbol{r},\boldsymbol{p},t)$, since the reflection of $\boldsymbol{v}$ is entirely equivalent to the reflection of $\boldsymbol{p}$ in the kinetic equation (see Equation (12) below). Thus the specular reflection conditions require for the distribution function $f(\boldsymbol{r},\boldsymbol{v},t)$ to satisfy

$$f(0,y,z,v_x,v_y,v_z,t)=f(0,y,z,-v_x,v_y,v_z,t) \quad \text{on} \quad x=0 \quad \text{plane}$$

$$f(a,y,z,v_x,v_y,v_z,t)=f(a,y,z,-v_x,v_y,v_z,t) \quad \text{on} \quad x=a \quad \text{plane}$$

$$f(x,0,z,v_x,v_y,v_z,t)=f(x,0,z,v_x,-v_y,v_z,t) \quad \text{on} \quad y=0 \quad \text{plane}$$

$$f(x,b,z,v_x,v_y,v_z,t)=f(x,b,z,v_x,-v_y,v_z,t) \quad \text{on} \quad y=b \quad \text{plane.}$$

In our duct-bounded plasma, the kinematic conditions on the four planes are satisfied by introducing extended electric field in the fashion:

$$E_x(-x,y,z)=-E_x(x,y,z), \quad E_x(2a-x,y,z)=-E_x(x,y,z) \tag{3}$$

$$E_y(x,-y,z)=-E_y(x,y,z), \quad E_y(x,2b-y,z)=-E_y(x,y,z) \tag{4}$$

This scheme is workable if $f_0(\boldsymbol{p})$, the zero order distribution function, is invariant with respect to the reflections $p_x\to -p_x$ and $p_y\to -p_y$, and clearly this reflectional property is satisfied by the moving Maxwellian to be introduced later. Equations (3) and (4) are the important conclusion of the above discussion that is valid in relativistic as well as in non-relativistic kinetic equations.

The function $E_x(x)$ as defined in Equation (3) is a periodic function of a piecewise continuous function of period "*a*" extending over the range $-\infty<x<\infty$ with discontinuity at $x=\pm 2na$ with a jump of $A_1$ (say) and with discontinuity at $x=\pm(2n-1)a$ with a jump of $A_2$ (say), where *n* is integer. The profile of the piecewise function $E_x(x)$ is plotted in the book by Lee [8]. The algebra involved in carrying out the Fourier transform of the piecewise discontinuous functions with aforementioned discontinuous jumps is quite laborious [7]. However, it turns out that, after all the algebraic hard work, the discontinuities that are present in the extended field components $E_x(x,y)$ and $E_y(x,y)$ in Equations (3) and (4) at the locations $x=\pm 2na$ and $x=\pm(2n-1)a$ and $y=\pm 2nb$ and $y=\pm(2n-1)b$ are mathematically (as well as physically) tantamount to placing fictitious surface charges at the corresponding jump locations in the form

$$\begin{aligned} S(x,y,z,t)=&A_1\sum_{n=0,1,2,\cdots}\delta(x\pm 2na)+A_2\sum_{n=1,2,\cdots}\delta(x\pm(2n-1)a)\\ &+B_1\sum_{n=0,1,2,\cdots}\delta(y\pm 2nb)+B_2\sum_{n=1,2,\cdots}\delta(y\pm(2n-1)b)\end{aligned} \tag{5}$$

This is the crucial part of improvement in this work as compared with the earlier work [7]. The surface charges are associated with the surface currents by sa-

tisfying the charge conservation equation

$$\frac{\partial S}{\partial t}+\nabla\cdot\hat{\boldsymbol{z}}\boldsymbol{J}_s=0 \tag{6}$$

Therefore, we can assume the presence of the fictitious surface currents

$$\boldsymbol{J}_s\left(\boldsymbol{k},\omega\right)=\hat{\boldsymbol{z}}\frac{k_z}{\omega}S\left(\boldsymbol{k},\omega\right) \tag{7}$$

The surface charges in Equation (5) and the surface currents in Equation (7) should be included in the Maxwell equations for our duct plasma wave analysis.

In this work, we consider a streaming plasma which moves along the axial direction ($z$-direction) with a relativistic speed $u$. In Section 2, we introduce the relativistic Vlasov equation. In Section 3, the boundary value problem satisfying the kinetic and electromagnetic boundary conditions are solved to find the dispersion relation of the duct surface wave. Section 4 furnishes discussions, and compares the dispersion relation with the recent work of Lee and Cho on non-relativistic duct flow [9].

## 2. Basic Equations

We begin with the relativistic equation of motion for electrons

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}=-e\left(\boldsymbol{E}+\frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\right) \tag{8}$$

where $\boldsymbol{p}$, the relativistic momentum, is

$$\boldsymbol{p}=m\gamma\boldsymbol{v} \tag{9}$$

with $m$ being the rest mass and $\boldsymbol{v}$ being the particle velocity, and

$$\gamma=\left(1-\frac{v^2}{c^2}\right)^{-1/2} \tag{10}$$

An easy way to write down the relativistic Vlasov equation is to note that the particle orbit as given by Equation (8) is the characteristics of the Vlasov equation. Indeed, the characteristic equation of the following equation is Equation (8):

$$\frac{\partial}{\partial t}f\left(\boldsymbol{r},\boldsymbol{p},t\right)+\boldsymbol{v}\cdot\frac{\partial f}{\partial\boldsymbol{r}}-e\left(\boldsymbol{E}\left(\boldsymbol{r},t\right)+\frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\right)\cdot\frac{\partial f}{\partial\boldsymbol{p}}=0 \tag{11}$$

Linearizing (11) gives

$$\frac{\partial}{\partial t}f\left(\boldsymbol{r},\boldsymbol{p},t\right)+\boldsymbol{v}\cdot\frac{\partial f}{\partial\boldsymbol{r}}-e\left(\boldsymbol{E}\left(\boldsymbol{r},t\right)+\frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\right)\cdot\frac{\partial f_0\left(\boldsymbol{p}\right)}{\partial\boldsymbol{p}}=0 \tag{12}$$

where $f_0\left(\boldsymbol{p}\right)$ is the zero order equilibrium distribution function which will be specified later. We also have the Maxwell equations for electrons. Ions are assumed to be stationary and only form the neutralizing background.

$$\nabla\times\boldsymbol{E}=-\frac{1}{c}\frac{\partial\boldsymbol{B}}{\partial t} \tag{13}$$

$$\nabla\times\boldsymbol{B}=\frac{4\pi}{c}\boldsymbol{J}+\frac{1}{c}\frac{\partial\boldsymbol{E}}{\partial t}+\hat{\boldsymbol{z}}J_s \tag{14}$$

where the plasma current $\boldsymbol{J}$ is

$$\boldsymbol{J}(\boldsymbol{r},t) = -\frac{eN}{m}\int \mathrm{d}^3 p\, f(\boldsymbol{r},\boldsymbol{p},t)\frac{\boldsymbol{p}}{\gamma} \tag{15}$$

with $N$ being the zero order equilibrium number density. The fictitious surface current $J_s$ corresponding to the surface charge S in Equation (5) takes the form

$$\begin{aligned} J_s = {} & A_1' \sum_{n=0,1,2,\cdots} \delta(x \pm 2na) + A_2' \sum_{n=1,2,\cdots} \delta\left(x \pm (2n-1)a\right) \\ & + B_1' \sum_{n=0,1,2,\cdots} \delta(y \pm 2nb) + B_2' \sum_{n=1,2,\cdots} \delta\left(y \pm (2n-1)b\right) \end{aligned} \tag{16}$$

$$\nabla \cdot \boldsymbol{E} = -4\pi e \int f \mathrm{d}^3 p + S(x,y,z) \tag{17}$$

$$\nabla \cdot \boldsymbol{B} = 0 \tag{18}$$

We Fourier transform all the dependent variables in the basic equations, including $f$, $\boldsymbol{E}$, and $\boldsymbol{B}$, and assume the form $\sim \exp(i\boldsymbol{k}\cdot\boldsymbol{r} - i\omega t)$. Then Equation (12) gives

$$f(\omega,\boldsymbol{k},\boldsymbol{p}) = i\frac{e}{m}\frac{1}{\omega - \boldsymbol{k}\cdot\boldsymbol{v}}\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\right)\cdot\frac{\partial f_0}{\partial \boldsymbol{p}} \tag{19}$$

where $\boldsymbol{v}$ reads as the function of $\boldsymbol{p}$ as given in Equation (9). Using Equation (19) in the Fourier transform of Equation (15) yields

$$J_l = -\frac{ie^2 N}{m}\int \mathrm{d}^3 p\, p_l \left[\frac{E_i + \frac{1}{c}e_{ijk}v_j B_k}{\gamma(\omega - \boldsymbol{k}\cdot\boldsymbol{v})}\right]\frac{\partial f_0}{\partial p_i} \tag{20}$$

where $l$ is the Cartesian index and $e_{ijk}$ is the Cartesian tensor called Levi-Civita symbol, and repeated indexes are summed over. We evaluate the current in Equation (20) for a cold streaming plasma whose zero order distribution function is

$$f_0(\boldsymbol{p}) = \delta(p_x)\delta(p_y)\delta(p_z - p_0) \tag{21}$$

where

$$p_0 = \gamma_0 m u,\ \gamma_0 = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{22}$$

which corresponds to the streaming velocity $\boldsymbol{v}_0 = \hat{\boldsymbol{z}}u$. Integrating by parts Equation (20) gives

$$J_l = \frac{ie^2 N}{m}\int \mathrm{d}^3 p\, f_0 \left(\delta_{il} + p_l \frac{\partial}{\partial p_i}\right)\left[\frac{E_i + \frac{1}{c}e_{ijk}v_j B_k}{\gamma(\omega - \boldsymbol{k}\cdot\boldsymbol{v})}\right] \tag{23}$$

The following relations are useful for evaluating the above integral:

$$\gamma(p) = \sqrt{1 + \frac{p^2}{m^2c^2}},\ \frac{\partial\gamma}{\partial p_j} = \frac{p_j}{\gamma m^2 c^2} \tag{24}$$

$$v_i = \frac{cp_i}{\sqrt{m^2c^2 + p^2}}, \quad \frac{\partial v_i}{\partial p_j} = \frac{1}{m\gamma}\left(\delta_{ij} - \frac{v_i v_j}{c^2}\right) \tag{25}$$

Using Equation (21) in Equation (23), we obtain after a long algebra,

$$\boldsymbol{J}(\boldsymbol{k},\omega) = \frac{ie^2N}{m\gamma_0\omega'}\left[\boldsymbol{E} + \frac{u}{c}\hat{\boldsymbol{z}}\times\boldsymbol{B} + \hat{\boldsymbol{z}}\left\{-\frac{u^2}{c^2}E_z + \frac{u}{\omega'}\left(\boldsymbol{k}\cdot\left(\boldsymbol{E} + \frac{u}{c}\hat{\boldsymbol{z}}\times\boldsymbol{B}\right) - \frac{u^2}{c^2}k_zE_z\right)\right\}\right] \tag{26}$$

where $\omega' = \omega - k_z u$ is Doppler-shifted frequency.

Using $\boldsymbol{B} = \frac{c}{\omega}\boldsymbol{k}\times\boldsymbol{E}$ and writing in components, we have

$$J_x = \frac{ie^2N}{m\omega\gamma_0}\left(E_x + \frac{uk_x}{\omega'}E_z\right) \tag{27}$$

$$J_y = \frac{ie^2N}{m\omega\gamma_0}\left(E_y + \frac{uk_y}{\omega'}E_z\right) \tag{28}$$

$$J_z = \frac{ie^2N}{m\omega\gamma_0}\left[\frac{u}{\omega'}\left(k_xE_x + k_yE_y\right) + \frac{E_z}{\omega'^2}\left(u^2\left(k_x^2 + k_y^2\right) + \frac{\omega^2}{\gamma_0^2}\right)\right] \tag{29}$$

The array formed by the coefficients of the electric field components is the conductivity tensor, as defined by $J_i = \sigma_{ij}E_j$.

Combining Equations (13) and (14) with the aid of Equations (27)-(29) yields a vector equation for $\boldsymbol{E}(\boldsymbol{k},\omega)$ in the form

$$\boldsymbol{k}\times\left(\boldsymbol{k}\times\boldsymbol{E}\right) + \frac{\omega^2}{c^2}\varepsilon_{ij}E_j = -\hat{\boldsymbol{z}}\frac{i\omega}{c}J_s \tag{30}$$

where

$$\varepsilon_{ij} = \delta_{ij} + \frac{4\pi i}{\omega}\sigma_{ij} = \left(1 - \frac{\omega_p^2}{\gamma_0\omega^2}\right)\delta_{ij} - \frac{\omega_p^2}{\gamma_0\omega^2}U_{ij} \tag{31}$$

is the dielectric tensor, and $\omega_p$ is the plasma frequency, and

$$U_{ij} = \begin{pmatrix} 0 & 0 & \frac{uk_x}{\omega'} \\ 0 & 0 & \frac{uk_y}{\omega'} \\ \frac{uk_x}{\omega'} & \frac{uk_y}{\omega'} & U_{zz} \end{pmatrix}$$

$$U_{zz} = \frac{u^2}{\omega'^2}\left(k^2 - \frac{\omega^2}{c^2}\right) + \frac{2k_zu}{\omega'}$$

Putting Equation (31) into Equation (30) gives a 3 by 3 matrix equation for $E_j$:

$$\left[\left(1 - n^2 - \frac{\omega_p^2}{\gamma_0\omega^2}\right)\delta_{ij} + \frac{k_ik_j}{k^2}n^2 - \frac{\omega_p^2}{\gamma_0\omega^2}U_{ij}\right]E_j = -\frac{ic}{\omega}\hat{\boldsymbol{z}}J_s \tag{32}$$

where $n^2 = c^2k^2/\omega^2$ is the refractive index, and $J_s$ is the Fourier transform of the fictitious surface currents in Equation (16):

$$J_s(\boldsymbol{k},\omega)=\delta(k_y)\left[A_1\Sigma_0 e^{\pm i2nak_x}+A_2\Sigma_1 e^{\pm i(2n-1)ak_x}\right]+\delta(k_x)\left[B_1\Sigma_0 e^{\pm i2nbk_y}+B_2\Sigma_1 e^{\pm i(2n-1)bk_y}\right] \tag{33}$$

where A's and B's may be functions of $k_z$ and $\omega$, and the double signs are summed over, and the notations $\Sigma_0$ and $\Sigma_1$ are the summations in Equation (16).

Equation (32) can be inverted after spending a considerable amount of time to solve for $\boldsymbol{E}(\boldsymbol{k},\omega)$:

$$E_x=-\frac{J_s}{\Delta}\frac{c^2k_x}{\omega^2}\left(k_z-\frac{u}{c^2}\frac{\omega_p^2}{\gamma_0\omega'}\right) \tag{34}$$

$$E_y=-\frac{J_s}{\Delta}\frac{c^2k_y}{\omega^2}\left(k_z-\frac{u}{c^2}\frac{\omega_p^2}{\gamma_0\omega'}\right) \tag{35}$$

$$E_z=\frac{J_s}{\Delta}\left(1-\frac{\omega_p^2}{\gamma_0\omega^2}-\frac{c^2k_z^2}{\omega^2}\right) \tag{36}$$

where

$$\Delta=\left(1-\frac{\omega_p^2}{\gamma_0\omega^2}-\frac{c^2k^2}{\omega^2}\right)\left[1-\frac{\omega_p^2}{\gamma_0\omega^2}-\frac{\omega_p^2}{\gamma_0\omega^2}\left(U_{zz}-\frac{u^2}{\omega'^2}\left(k_x^2+k_y^2\right)\right)\right] \tag{37}$$

After further algebra, $\Delta$ becomes

$$\Delta=\left(1-\frac{\omega_p^2}{\gamma_0\omega^2}-\frac{c^2k^2}{\omega^2}\right)\left(1-\frac{\omega_p^2}{\gamma_0^3\omega'^2}\right) \tag{38}$$

Also we obtain

$$B_x=\frac{c}{\omega}\frac{J_s}{\Delta}k_y\left(1-\frac{\omega_p^2}{\gamma_0\omega\omega'}\right) \tag{39}$$

$$B_y=-\frac{c}{\omega}\frac{J_s}{\Delta}k_x\left(1-\frac{\omega_p^2}{\gamma_0\omega\omega'}\right) \tag{40}$$

In addition, we put $B_z=0$ since we investigate the transverse magnetic mode.

## 3. Boundary Equations

In order to apply the boundary conditions, the electric and magnetic field components in the Fourier $\boldsymbol{k}$ space should be inverted to the fields in the ordinary $\boldsymbol{r}$ space by performing $\int_{-\infty}^{\infty}dk_x\int_{-\infty}^{\infty}dk_y\,e^{ik_xx+ik_yy}(\cdots)$. The integrals involve infinite series through the surface charge $J_s$, but the infinite series are nicely summed at the particular positions corresponding to $x=0,a$ and $y=0,b$. Thus, we apply the boundary conditions along the two infinite lines: $(x,y,z)=(0,0,z)$ and $(a,b,z)$ with $-\infty<z<\infty$. The two lines correspond to the two seams of the duct which are diagonally opposite.

When the inversion integrals are performed, the following formulas are useful, which can be verified by a simple change of variable, as is shown in earlier work

[7]. We have integrals of the type in the inversion integrals

$$Q(x)=\int_{-\infty}^{\infty}\mathrm{d}k_x k_x \Phi(k)\mathrm{e}^{ik_x x}\left[A_1\Sigma_0\mathrm{e}^{\pm i2nak_x}+A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\right] \tag{41}$$

where $\Phi(k)$ is an even function of $k_x$. Then, we have

$$Q(0)=A_1\int_{-\infty}^{\infty}\mathrm{d}k_x k_x \Phi(k_x) \tag{42}$$

$$Q(a)=-A_2\int_{-\infty}^{\infty}\mathrm{d}k_x k_x \Phi(k_x) \tag{43}$$

We also have integrals of the type

$$R(x)=\int_{-\infty}^{\infty}\mathrm{d}k_x \Phi(k)\mathrm{e}^{ik_x x}\left[A_1\Sigma_0\mathrm{e}^{\pm i2nak_x}+A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\right] \tag{44}$$

Then, we have

$$R(0)=2\int_{-\infty}^{\infty}\mathrm{d}k_x \Phi(k_x)\left(A_1S_1+A_2S_2\right) \tag{45}$$

$$R(a)=2\int_{-\infty}^{\infty}\mathrm{d}k_x \Phi(k_x)\left(A_1S_2+A_2S_1\right) \tag{46}$$

where

$$S_1=\frac{1}{2}+\mathrm{e}^{2iak_x}+\mathrm{e}^{4iak_x}+\cdots \tag{47}$$

$$S_2=\mathrm{e}^{iak_x}+\mathrm{e}^{3iak_x}+\cdots \tag{48}$$

Formulas in Equations (42), (43), (45), and (46) are useful for evaluating the integrals. Let us evaluate:

$$\begin{aligned}
&E_x(0,0,z)\\
&=-\frac{c^2}{\omega^2}\left(k_z-\frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)\int_{-\infty}^{\infty}\mathrm{d}k_x\int_{-\infty}^{\infty}\mathrm{d}k_y\frac{k_x}{\Delta}\Big[\delta(k_y)\Big(A_1\Sigma_0\mathrm{e}^{\pm i2nak_x}\\
&\quad+A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\Big)+\delta(k_x)\Big(B_1\Sigma_0\mathrm{e}^{\pm i2nbk_y}+B_2\Sigma_1\mathrm{e}^{\pm i(2n-1)bk_y}\Big)\Big]\\
&=-\frac{c^2}{\omega^2}\left(k_z-\frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)\int_{-\infty}^{\infty}\mathrm{d}k_x\frac{k_x}{\Delta}\Big(A_1\Sigma_0\mathrm{e}^{\pm i2nak_x}+A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\Big)\\
&=-\frac{c^2}{\omega^2}\left(k_z-\frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)A_1\int_{-\infty}^{\infty}\mathrm{d}k_x\frac{k_x}{\Delta}
\end{aligned} \tag{49}$$

where we used Equation (42). In the last (also in the later) integral, $k^2$ hidden in $\Delta$ is $k^2=k_z^2+k_x^2$. We have

$$\begin{aligned}
&E_x(a,b,z)\\
&=-\frac{c^2}{\omega^2}\left(k_z-\frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)\int_{-\infty}^{\infty}\mathrm{d}k_x\mathrm{e}^{ik_x a}\int_{-\infty}^{\infty}\mathrm{d}k_y\mathrm{e}^{ik_y b}\frac{k_x}{\Delta}\Big[\delta(k_y)\Big(A_1\Sigma_0\mathrm{e}^{\pm i2nak_x}\\
&\quad+A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\Big)+\delta(k_x)\Big(B_1\Sigma_0\mathrm{e}^{\pm i2nbk_y}+B_2\Sigma_1\mathrm{e}^{\pm i(2n-1)bk_y}\Big)\Big]\\
&=-\frac{c^2}{\omega^2}\left(k_z-\frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)\int_{-\infty}^{\infty}\mathrm{d}k_x\mathrm{e}^{ik_x a}\frac{k_x}{\Delta}\Big(A_1\Sigma_0{}^{\pm i2nak_x}+A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\Big)\\
&=-\frac{c^2}{\omega^2}\left(k_z-\frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)(-A_2)\int_{-\infty}^{\infty}\mathrm{d}k_x\frac{k_x}{\Delta}
\end{aligned} \tag{50}$$

where we used Equation (43). Analogous integrations yield

$$E_y(0,0,z) = -\frac{c^2}{\omega^2}\left(k_z - \frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)B_1\int_{-\infty}^{\infty}\mathrm{d}k_y\,\frac{k_y}{\Delta} \tag{51}$$

$$E_y(a,b,z) = -\frac{c^2}{\omega^2}\left(k_z - \frac{u\omega_p^2}{c^2\omega'\gamma_0}\right)(-B_2)\int_{-\infty}^{\infty}\mathrm{d}k_y\,\frac{k_y}{\Delta} \tag{52}$$

In the above (also in the later) $\int \mathrm{d}k_y$ integral, $k^2$ hidden in $\Delta$ is $k^2 = k_z^2 + k_y^2$.

$$B_x(0,0,z) = \frac{c}{\omega}\left(1 - \frac{\omega_p^2}{\omega\omega'\gamma_0}\right)B_1\int_{-\infty}^{\infty}\mathrm{d}k_y\,\frac{k_y}{\Delta} \tag{53}$$

$$B_x(a,b,z) = \frac{c}{\omega}\left(1 - \frac{\omega_p^2}{\omega\omega'\gamma_0}\right)(-B_2)\int_{-\infty}^{\infty}\mathrm{d}k_y\,\frac{k_y}{\Delta} \tag{54}$$

$$B_y(0,0,z) = -\frac{c}{\omega}\left(1 - \frac{\omega_p^2}{\omega\omega'\gamma_0}\right)A_1\int_{-\infty}^{\infty}\mathrm{d}k_x\,\frac{k_x}{\Delta} \tag{55}$$

$$B_y(a,b,z) = -\frac{c}{\omega}\left(1 - \frac{\omega_p^2}{\omega\omega'\gamma_0}\right)(-A_2)\int_{-\infty}^{\infty}\mathrm{d}k_x\,\frac{k_x}{\Delta} \tag{56}$$

We encounter with different type of integral in

$$E_z(0,0,z) = \left(1 - \frac{\omega_p^2}{\gamma_0\omega^2} - \frac{c^2k_z^2}{\omega^2}\right)\int_{-\infty}^{\infty}\mathrm{d}k_x\int_{-\infty}^{\infty}\mathrm{d}k_y\,\frac{1}{\Delta}\Big[\delta(k_y)\left(A_1\Sigma_0\,\mathrm{e}^{\pm i2nak_x} + A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\right) + \delta(k_x)\left(B_1\Sigma_0\mathrm{e}^{\pm i2nbk_y} + B_2\Sigma_1\mathrm{e}^{\pm i(2n-1)bk_y}\right)\Big]$$

which becomes

$$E_z(0,0,z) = \left(1 - \frac{\omega_p^2}{\gamma_0\omega^2} - \frac{c^2k_z^2}{\omega^2}\right)\Big[\int_{-\infty}^{\infty}\frac{\mathrm{d}k_x}{\Delta}\left(A_1\Sigma_0\mathrm{e}^{\pm i2nak_x} + A_2\Sigma_1\mathrm{e}^{\pm i(2n-1)ak_x}\right) + \int_{-\infty}^{\infty}\frac{\mathrm{d}k_y}{\Delta}\left(B_1\Sigma_0\mathrm{e}^{\pm i2nbk_y} + B_2\Sigma_1\mathrm{e}^{\pm i(2n-1)bk_y}\right)\Big]$$

which we write in the form

$$E_z(0,0,z) = 2\left(1 - \frac{\omega_p^2}{\gamma_0\omega^2} - \frac{c^2k_z^2}{\omega^2}\right)\Big[\int_{-\infty}^{\infty}\frac{\mathrm{d}k_x}{\Delta}\left(A_1S_1(ak_x) + A_2S_2(ak_x)\right) + \int_{-\infty}^{\infty}\frac{\mathrm{d}k_y}{\Delta}\left(B_1S_1(bk_y) + B_2S_2(bk_y)\right)\Big] \tag{57}$$

where we used Equation (45), and

$$S_1(\xi) = \frac{1}{2} + \mathrm{e}^{2i\xi} + \mathrm{e}^{4i\xi} + \cdots, \quad S_2(\xi) = \mathrm{e}^{i\xi} + \mathrm{e}^{3i\xi} + \cdots \tag{58}$$

Analogously, we obtain

$$E_z(a,b,z) = 2\left(1 - \frac{\omega_p^2}{\gamma_0\omega^2} - \frac{c^2k_z^2}{\omega^2}\right)\Big[\int_{-\infty}^{\infty}\frac{\mathrm{d}k_x}{\Delta}\left(A_1S_2(ak_x) + A_2S_1(ak_x)\right) + \int_{-\infty}^{\infty}\frac{\mathrm{d}k_y}{\Delta}\left(B_1S_2(bk_y) + B_2S_1(bk_y)\right)\Big] \tag{59}$$

where we used Equation (46).

- **Vacuum solution**.

Vacuum solutions should solve

$$\left(\nabla^2 + \frac{\omega^2}{c^2}\right)\boldsymbol{B} = 0 \tag{60}$$

and

$$\boldsymbol{E} = \frac{ic}{\omega}\nabla \times \boldsymbol{B} \tag{61}$$

Equation (60) is solved by

$$\boldsymbol{B} \sim \mathrm{e}^{ik_z z}\mathrm{e}^{\pm k_x x}\mathrm{e}^{\pm k_y y} \tag{62}$$

with constraint $k_x^2 + k_y^2 = k_z^2 - \frac{\omega^2}{c^2} \equiv \lambda^2$ and $\nabla \cdot \boldsymbol{B} = 0$. Furthermore, we assume $B_z = 0$ since we consider the TM (transverse magnetic) mode.

The vacuum regions corresponding to (or exterior to) the lines $(0,0,z)$ and $(a,b,z)$, which we designate as (*i*) and (*ii*), respectively, are:

Vacuum region (*i*) $x < 0, y <$ , where we have

$$B_x^v(i) = H_x \mathrm{e}^{ik_z z}\mathrm{e}^{k_x x}\mathrm{e}^{k_y y} \tag{63}$$

$$B_y^v(i) = H_y \mathrm{e}^{ik_z z}\mathrm{e}^{k_x x}\mathrm{e}^{k_y y} \tag{64}$$

$$k_x H_x + k_y H_y = 0 \tag{65}$$

$$E_z^v(i) = \frac{ic}{\omega}\left(H_y k_x - H_x k_y\right)\mathrm{e}^{ik_z z}\mathrm{e}^{k_x x}\mathrm{e}^{k_y y} \tag{66}$$

$$E_x^v(i) = \frac{c}{\omega}k_z H_y \mathrm{e}^{ik_z z}\mathrm{e}^{k_x x}\mathrm{e}^{k_y y} \tag{67}$$

$$E_y^v(i) = -\frac{c}{\omega}k_z H_x \mathrm{e}^{ik_z z}\mathrm{e}^{k_x x}\mathrm{e}^{k_y y} \tag{68}$$

Vacuum region (*ii*) $x > a, y > b$, where

$$B_x^v(ii) = G_x \mathrm{e}^{ik_z z}\mathrm{e}^{-k_x x}\mathrm{e}^{-k_y y} \tag{69}$$

$$B_y^v(ii) = G_y \mathrm{e}^{ik_z z}\mathrm{e}^{-k_x x}\mathrm{e}^{-k_y y} \tag{70}$$

$$k_x G_x + k_y G_y = 0 \tag{71}$$

$$E_z^v(ii) = \frac{ic}{\omega}\left(-G_y k_x + G_x k_y\right)\mathrm{e}^{ik_z z}\mathrm{e}^{-k_x x}\mathrm{e}^{-k_y y} \tag{72}$$

$$E_x^v(ii) = \frac{c}{\omega}k_z G_y \mathrm{e}^{ik_z z}\mathrm{e}^{-k_x x}\mathrm{e}^{-k_y y} \tag{73}$$

$$E_y^v(ii) = -\frac{c}{\omega}k_z G_x \mathrm{e}^{ik_z z}\mathrm{e}^{-k_x x}\mathrm{e}^{-k_y y} \tag{74}$$

Putting $(x,y) = (0,0)$ or $(a,b)$ in the above equations gives the vacuum side values of the relevant quantities.

## 4. Dispersion Relation

We enforce the following boundary conditions to connect the plasma and the vacuum fields: $[E_z]=0$, $[B_y]=\frac{u}{c}[E_x]$, $[B_x]=-\frac{u}{c}[E_y]$.

Along line $(0,0,z)$

$[E_z]=0$ gives, per Equations (57) and (66),

$$\left(1-\frac{\omega_p^2}{\gamma_0\omega^2}-\frac{c^2k_z^2}{\omega^2}\right)\left(A_1I_1+A_2I_2+B_1J_1+B_2J_2\right)=\frac{ic}{\omega}\left(H_yk_x-H_xk_y\right) \tag{75}$$

where

$$I_i=2\int_{-\infty}^{\infty}\frac{\mathrm{d}k_x}{\Delta}S_i\left(ak_x\right),\ J_i=2\int_{-\infty}^{\infty}\frac{\mathrm{d}k_x}{\Delta}S_i\left(bk_y\right),\ (i=1,2) \tag{76}$$

$[B_y]=\frac{u}{c}[E_x]$ gives

$$A_1Q\left(1-\frac{\omega_p^2}{\gamma_0^3\omega'^2}\right)+\frac{\omega}{c}H_y=0 \tag{77}$$

where

$$Q=\int_{-\infty}^{\infty}\mathrm{d}k_x\frac{k_x}{\Delta}=\int_{-\infty}^{\infty}\mathrm{d}k_y\frac{k_y}{\Delta} \tag{78}$$

$[B_x]=-\frac{u}{c}[E_y]$ gives

$$B_1Q\left(1-\frac{\omega_p^2}{\gamma_0^3\omega'^2}\right)-\frac{\omega}{c}H_x=0 \tag{79}$$

Along line $(a,b,z)$

$[E_z]=0$ gives

$$\begin{aligned}&\left(1-\frac{\omega_p^2}{\gamma_0\omega^2}-\frac{c^2k_z^2}{\omega^2}\right)\left(A_1I_2+A_2I_1+B_1J_2+B_2J_1\right)\\&=\frac{ic}{\omega}\left(-G_yk_x+G_xk_y\right)\mathrm{e}^{-k_xa}\mathrm{e}^{-k_yb}\end{aligned} \tag{80}$$

$[B_y]=\frac{u}{c}[E_x]$ gives

$$A_2Q\left(1-\frac{\omega_p^2}{\gamma_0^3\omega'^2}\right)-\frac{\omega}{c}G_y\mathrm{e}^{-k_xa}\mathrm{e}^{-k_yb}=0 \tag{81}$$

$[B_y]=-\frac{u}{c}[E_x]$ gives

$$B_2Q\left(1-\frac{\omega_p^2}{\gamma_0^3\omega'^2}\right)+\frac{\omega}{c}G_x\mathrm{e}^{-k_xa}\mathrm{e}^{-k_yb}=0 \tag{82}$$

In addition, we have, per $\nabla\cdot\boldsymbol{B}=0$ and $B_z=0$,

$$k_xH_x+k_yH_y=0 \tag{83}$$

$$k_x G_x + k_y G_y = 0 \tag{84}$$

Thus, we have 8 equations for 8 unknowns; $A_1, A_2, B_1, B_2, H_x, H_y, G_x, G_y$. Eliminating $H_x, H_y, G_x, G_y$ gives

$$A_1\left(I_1 + i\frac{c^2}{\omega^2}\frac{\eta}{\xi}k_x Q\right) + A_2 I_2 + B_1\left(J_1 + i\frac{c^2}{\omega^2}\frac{\eta}{\xi}k_y Q\right) + B_2 J_2 = 0 \tag{85}$$

$$A_1 I_2 + A_2\left(I_1 + i\frac{c^2}{\omega^2}\frac{\eta}{\xi}k_x Q\right) + B_1 J_2 + B_2\left(J_1 + i\frac{c^2}{\omega^2}\frac{\eta}{\xi}k_y Q\right) = 0 \tag{86}$$

$$k_y A_1 = k_x B_1 \tag{87}$$

$$k_y A_2 = k_x B_2 \tag{88}$$

where

$$\xi = 1 - \frac{\omega_p^2}{\gamma_0 \omega^2} - \frac{c^2 k_z^2}{\omega^2}, \quad \eta = 1 - \frac{\omega_p^2}{\gamma_0^3 \omega'^2} \tag{89}$$

Eliminating $B_1$ and $B_2$ gives

$$A_1\left[k_x I_1 + k_y J_1 + \frac{ic^2}{\omega^2}\frac{\eta}{\xi}\left(k_x^2 + k_y^2\right)Q\right] + A_2\left(k_x I_2 + k_y J_2\right) = \tag{90}$$

$$A_1\left(k_x I_2 + k_y J_2\right) + A_2\left[k_x I_1 + k_y J_1 + \frac{ic^2}{\omega^2}\frac{\eta}{\xi}\left(k_x^2 + k_y^2\right)Q\right] = 0 \tag{91}$$

The above two equations for $A_1$ and $A_2$ yield the dispersion relation in the form

$$k_x \int_{-\infty}^{\infty} \frac{\mathrm{d}k_x}{\Delta} \frac{1 \pm \mathrm{e}^{iak_x}}{1 \mp \mathrm{e}^{iak_x}} + k_y \int_{-\infty}^{\infty} \frac{\mathrm{d}k_y}{\Delta} \frac{1 \pm \mathrm{e}^{ibk_y}}{1 \mp \mathrm{e}^{ibk_y}} + i\lambda^2 \frac{c^2}{\omega^2}\frac{\eta}{\xi}\int_{-\infty}^{\infty} k_x \frac{\mathrm{d}k_x}{\Delta} = 0 \tag{92}$$

where we used

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}k_x}{\Delta}\left[S_1(ak_x) \pm S_2(ak_x)\right] = \frac{1}{2}\int_{-\infty}^{\infty} \frac{\mathrm{d}k_x}{\Delta} \frac{1 \pm \mathrm{e}^{iak_x}}{1 \mp \mathrm{e}^{iak_x}} \tag{93}$$

In regard to the Fourier variables $k_x$ and $k_y$ outside the integrals, we imposed the constraint $k_x^2 + k_y^2 = k_z^2 - \omega^2/c^2 \equiv \lambda^2$. Therefore it is convenient to transform

$$k_x = \frac{b\lambda}{\sqrt{a^2 + b^2}}, \quad k_y = \frac{a\lambda}{\sqrt{a^2 + b^2}} \tag{94}$$

[$k_x, k_y$ inside the integrals are dummy and let them be there as they are.] The transform in Equation (94) satisfies the constraint and the relation $ak_x = bk_y$. In fact it can be derived from the latter and the constraint. Then the dispersion relation takes the form

$$\begin{aligned} &\frac{b}{\sqrt{a^2 + b^2}} \int_{-\infty}^{\infty} \frac{\mathrm{d}k_x}{\Delta} \frac{1 \pm \mathrm{e}^{iak_x}}{1 \mp \mathrm{e}^{iak_x}} + \frac{a}{\sqrt{a^2 + b^2}} \int_{-\infty}^{\infty} \frac{\mathrm{d}k_y}{\Delta} \frac{1 \pm \mathrm{e}^{ibk_y}}{1 \mp \mathrm{e}^{ibk_y}} \\ &+ i\lambda \frac{c^2}{\omega^2}\frac{\eta}{\xi}\int_{-\infty}^{\infty} k_x \frac{\mathrm{d}k_x}{\Delta} = 0 \end{aligned} \tag{95}$$

If either $a$ or $b \to \infty$, we recover the slab dispersion relation in the

non-relativistic limit [7].

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}k_x}{\Delta} \frac{1 \pm \mathrm{e}^{iak_x}}{1 \mp \mathrm{e}^{iak_x}} + i\lambda \frac{c^2}{\omega^2} \frac{\eta}{\xi} \int_{-\infty}^{\infty} k_x \frac{\mathrm{d}k_x}{\Delta} = 0 \tag{96}$$

where $\eta$ and $\xi$ are given by Equation (89) with $\gamma_0 = 1$.

It is recalled that $k^2$ hidden in $\Delta$ is: $k^2 = k_z^2 + k_x^2$ in $\int \mathrm{d}k_x$-integral and $k^2 = k_z^2 + k_y^2$ in $\int \mathrm{d}k_y$-integral. Thus, let us change the integration variables, both $k_x$ and $k_y$, in Equation (95) to $\kappa$:

$$\begin{aligned} &\frac{b}{\sqrt{a^2+b^2}} \int_{-\infty}^{\infty} \frac{\mathrm{d}\kappa}{\Delta} \frac{1 \pm \mathrm{e}^{ia\kappa}}{1 \mp \mathrm{e}^{ia\kappa}} + \frac{a}{\sqrt{a^2+b^2}} \int_{-\infty}^{\infty} \frac{\mathrm{d}\kappa}{\Delta} \frac{1 \pm \mathrm{e}^{ib\kappa}}{1 \mp \mathrm{e}^{ib\kappa}} \\ &+ i\lambda \frac{c^2}{\omega^2} \frac{\eta}{\xi} \int_{-\infty}^{\infty} \kappa \frac{\mathrm{d}\kappa}{\Delta} = 0 \end{aligned} \tag{97}$$

where $k^2 = k_z^2 + \kappa^2$. In regard to the double signs in Equation (97), the upper (lower) signs correspond to the symmetric (anti-symmetric) mode which also occurs in a slab plasma. For a square duct ($a = b$), Equation (97) reduces to the form identical with the slab dispersion equation Equation (96), except for factor $\sqrt{2}$. This reduction is due to the *x*-*y* symmetry. To recover slab dispersion relation from Equation (92), we take $k_y \to 0$, $b \to \infty$, and put $k_x = \lambda$. We can take $k_y \to 0$ since the y-direction has a translational invariance in a slab.

The duct dispersion relation in Equation (97) can be contour-integrated for a cold plasma, giving

$$\frac{b\Gamma}{\sqrt{a^2+b^2}} \tanh\frac{a\Gamma}{2} + \frac{a\Gamma}{\sqrt{a^2+b^2}} \tanh\frac{b\Gamma}{2} + \sqrt{k_z^2 - \frac{\omega^2}{c^2}} \left(1 - \frac{\omega_p^2}{\gamma_0^3 \omega'^2}\right) = 0 \tag{98}$$

where $\Gamma = \sqrt{k_z^2 - \dfrac{\omega^2 - \omega_p^2/\gamma_0}{c^2}}$. For the anti-symmetric mode, tanh-function above is replaced by coth-function.

## 5. Discussion

In a bounded plasma, one way of solving Vlasov equation by satisfying the specular reflection condition is to extend the plasma electric field in the manner of Equation (3). The job of Fourier transforming such as piecewise continuous periodic function, extending to infinity, is laborious. In this work, we present an alternative way of avoiding the hard algebra by placing sheets of fictitious surface charges at the location of discontinuities of the electric field. The magnitudes of the surface charges are undetermined constants, but they can be determined through the connection formula with the vacuum side field—resulting in the dispersion relation of the surface wave. This method enables one to deal with semi-infinite, slab, and duct plasmas in a common work-frame. Taking $b \to \infty$ in Equation (98) gives

$$\Gamma \tanh\frac{a\Gamma}{2} + \sqrt{k_z^2 - \frac{\omega^2}{c^2}} \left(1 - \frac{\omega_p^2}{\gamma_0^3 \omega'^2}\right) = 0 \tag{99}$$

which is the slab ( $0 < x < a$ ) dispersion relation. Taking $a \to \infty$ in Equation (99) gives the semi-infinite plasma dispersion relation

$$\Gamma + \sqrt{k_z^2 - \frac{\omega^2}{c^2}}\left(1 - \frac{\omega_p^2}{\gamma_0^3 \omega'^2}\right) = 0 \tag{100}$$

Equation (100) agrees with the semi-infinite dispersion relation obtained by Lee [3]. If $u = 0$, Equation (99) agrees with the slab dispersion relation obtained from the fluid theory worked out by Gradov and Stenflo [10].

For a square duct, putting $a = b$ in Equation (98) yields

$$\sqrt{2}\Gamma \tanh\frac{a\Gamma}{2} + \sqrt{k_z^2 - \frac{\omega^2}{c^2}}\left(1 - \frac{\omega_p^2}{\gamma_0^3 \omega'^2}\right) = 0 \tag{101}$$

which is similar to the slab dispersion relation. This is because the complete symmetry between $x$ and $y$ coordinates makes the three-dimensional problem a two-dimensional problem practically.

The Doppler-shifted frequency $\omega'$ appearing in Equation (98) represents the streaming effect in the lowest order of $\frac{u}{c}$. The higher order effect enters through the relativistic factor $\gamma_0$ and $U_{zz}$ in Equation (31). The relativistic effect manifests itself through the attenuation constant $\Gamma$ per $\frac{\omega_p^2}{\gamma_0}$ and in the dispersion relation per $\frac{\omega_p^2}{\gamma_0^3}$. Recently Lee and Cho [9] investigated surface waves in a non-relativistically streaming Vlasov plasma in a duct. Their result is identical with Equation (98) upon putting $\gamma_0 = 1$.

A visual understanding of the extended electric field may be grasped by plotting the extended function [8]. A useful reference for relativistic Vlasov equation is, among others, Momtgomery and Tidman [11], in which the velocity-version of Vlasov equation (in contrast to the momentum-version employed in this work) is presented in detail. This work may find applications in laboratory or astrophysical situation where electromagnetic waves propagate through certain channels. This work might be useful for analysis of a proton beam travelling in a duct.

## Acknowledgements

This research was supported by National Cancer Center Grant (NCC-2110370), Korea.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Lee, H.J. and Cho, S.H. (1997) *Journal of Plasma Physics*, **58**, 409-419. https://doi.org/10.1017/S002237789700603X

[2] Lee, H.J. and Cho, S.H. (1999) *Journal of Plasma Physics*, **61**, 173-174. https://doi.org/10.1017/S002237789800720X

[3] Lee, H.J. (2005) *Physics of Plasmas*, **12**, 094701-1-4. https://doi.org/10.1063/1.1878832

[4] Landau, L.D. (1946) *Journal of Physics* (*USSR*), **10**, 25.

[5] Barr, H.C. and Boyd, T.J.M. (1972) *Journal of Physics A*: *General Physics*, **5**, 1108-1118. https://doi.org/10.1088/0305-4470/5/7/019

[6] Alexandrov, A.F., Bogdankevich, L.S. and Rukhadze, A.A. (1984) Principles of Plasma Electrodynamics. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 159. https://doi.org/10.1007/978-3-642-69247-5

[7] Lee, H.J. and Lim, Y.K. (2007) *Journal of the Korean Physical Society*, **50**, 1056-1061. https://doi.org/10.3938/jkps.50.1056

[8] Lee, H.J. (2019) Fundamentals of Theoretical Plasma Physics: Mathematical Description of Plasma Waves. World Scientific, Singapore, 551. https://doi.org/10.1142/11168

[9] Lee, H.J. and Cho, S.-H. (2022) *Journal of Plasma Physics*, **88**, Article ID: 905880305. https://doi.org/10.1017/S0022377822000332

[10] Gradov, O.M., *et al.* (1983) *Physics Reports*, **94**, 111-137. https://doi.org/10.1016/0370-1573(83)90004-2

[11] Montgomery, D.C. and Tidman, D.A. (1964) Plasma Kinetic Theory. McGraw-Hill, New York.

Journal of Modern Physics, 2022, 13, 1080-1092
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Circular Scale of Time as a Guide of the Schrödinger's Perturbation Theory

**S. Olszewski**

Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland
Email: solszewski@ichf.edu.pl

**How to cite this paper:** Olszewski, S. (2022) Circular Scale of Time as a Guide of the Schrödinger's Perturbation Theory. *Journal of Modern Physics*, **13**, 1080-1092.
https://doi.org/10.4236/jmp.2022.137061

**Received:** May 25, 2022
**Accepted:** July 15, 2022
**Published:** July 18, 2022



## Abstract

The paper is a kind of a review which considers an investigation of the scale of time suggested by an application of the Schrödinger perturbation method, especially when the perturbation of a non-degenerate quantum state is examined. In fact the method was applied in numerous cases—also by Schrödinger himself—without any use of the notion of time. Simultaneously, because of the development of computers, their use in solving the perturbation problems gradually decreased. However, the point of importance in the paper became the time. We demonstrate that collisions of a quantum system with the perturbation potential can be arranged along a circular scale of time whose properties provide us precisely with the energy terms obtained by the Schrödinger perturbation theory. This validity of results is checked till the perturbation order $N = 7$.

## Keywords

Scale of Time, Schrödinger's Perturbation Theory, Non-Degenerate Quantum State

## 1. Introduction. Different Kinds of Approach to the Time Parameter in the Everyday Life and Science

Evidently the time was an important parameter in the human existence from its very beginning.

Duration and repetition of the days, nights, seasons, years became a well-known observation of everybody. This situation did not change in course of centuries. Simultaneously we had, in general, a strict qualitative distinction between the intervals of time and space.

The space intervals were easy to manipulate in their arrangement, both in imagination and practice: there was no difficulty to have or put any such interval

in an arbitrary position or direction chosen by the observer. A totally different property concerned the intervals of time: they had always a definite property of a future object, or a past object, or an object being actually present in our interest.

Nevertheless the science, especially mechanics, could be developed in spite of a difficulty concerning the actual "historical" position of an interval of time. In effect the time interval entering the mechanical process could be considered independently from its "history" for, in many occasions, the mechanics could be liberated from its historical background associated with time.

A special point which made the sense of mechanical laws questionable was connected with an examination of the physical laws concerning the whole mechanical systems. In this case the main result became that a mechanical system, having a constant velocity, should not obey several kinetic laws other than those obtained for a system at rest; see e.g. [1]. This is usually presented by a requirement that the Galilean transformation laws for the mechanical parameters have to be valid. But the development done in physics in the 19th century led to conclusion that the laws of the Maxwell electrodynamics should be equally valid in a moving system as well as they are satisfied for a system at rest.

A well-known consequence of that conclusion was the replacement of the Galilean transformation of the mechanical parameters by the Lorentz one. In fact, the Galilean transformation keeps its good accuracy solely when the speed of the moving mechanical system remains low in comparison with the speed of light $c$—the effect which holds in the most part of situations met in the everyday life.

## 2. Present Approach to the Problem of Time and Its Scientific Position

A competition between the Galilean and Lorentz transformations done by the Lorentz formula, presented a well-known subject of the special relativistic theory. This theory applies the joint metrics of the time interval

$$\mathrm{d}t \tag{1}$$

and space intervals

$$\mathrm{d}x, \mathrm{d}y, \mathrm{d}z \tag{2}$$

by combining them into the formula

$$\left(\mathrm{d}s\right)^2 = c^2\left(\mathrm{d}t\right)^2 - \left(\mathrm{d}x\right)^2 - \left(\mathrm{d}y\right)^2 - \left(\mathrm{d}z\right)^2 \tag{3}$$

representing the square of a small distant $\mathrm{d}s$ of the moving body. In the general theory of relativity the metrics (3) is replaced by a more complicated one [2]

$$-\left(\mathrm{d}s\right)^2 = g_{ik}\mathrm{d}x_i\,\mathrm{d}x_k. \tag{4}$$

in which summation does apply over the parameters $i$ and $k$. Usually the space coordinates are $x_1$, $x_2$ and $x_3$ and the time coordinate is denoted by $x_0$. In general the terms in (4) being

$$g_{ik} \tag{5}$$

are some functions of $x_i$ and $x_k$. A scientific advantage of the metrics due to

(4) and (5) over the metrics (3) is that (4) and (5) can take into account several special physical effects, like the gravitational interaction between mass and light confirmed next by the observation.

An outline of the ideas and formulas given above concerns mainly the classical physics. They allow us, however, to present the role of time in a different problem, referred mainly to the quantum theory. This theory, began by the Planck's treatment of the oscillators entering the black-body ensemble, allowed him to discover the oscillator quanta of energy, as well as the roles of the oscillator frequency and the constant carrying the Planck's name.

The next large step towards quanta was connected with a partly quantum and partly classical approach to the hydrogen atom developed by Bohr; see [3], Vol. 1. Because of its very good agreement with the observed data, the model was considered as practically perfect in calculating the light frequencies connected with the electron transitions in the atom. But next the applications of the quantum theory occurred rather limited because of the difficulty connected with a treatment of the many-electron systems present in the non-hydrogen atoms. This difficulty was successfully defeated by Schrödinger—and his successors—in the wave-mechanical approach to the electron structure of the atoms; see e.g. [3], Vol. 2.

## 3. Schrödinger's Quantum Problems and Simplification of Their Solutions

The main idea of Schrödinger was to follow the de Broglie concept and consider the electron as a wave-like particle of matter.

Then a corresponding wave-like equation can be built up and next solved. One side of the equation is a sum of the kinetic energy operator of one or many electrons presented in a system, and the next term in the sum is the potential energy operator which takes different particle interactions necessary to be considered into account. Another side of the Schrödinger equation is given by a product of the energy constant $E$ multiplied by the electron wave function $\psi$. In effect we obtain the eigenequation for $E$ and $\psi$. In general its solution represents a complicated mathematical task—only for very simple physical systems the equation can be rather readily solved.

The Schrödinger equation—on its one side—is a sum of the kinetic and potential energy operators $\hat{E}_{\text{kin}}$ and $\hat{E}_{\text{pot}}$, viz.

$$\hat{H} = \hat{E}_{\text{kin}} + \hat{E}_{\text{pot}} \tag{6}$$

called the Hamiltonian—or energy—operator. This operator is acting on the wave function $\psi$, so

$$\hat{H}\psi = E\psi, \tag{7}$$

is giving the Schrödinger equation. The right-hand side of the Equation in (7) represents a product of the eigenenergy $E$, considered as a constant number, and $\psi$. The effect of solution of (7) is usually a discrete set of values of $E$ and dis-

crete set of functions $\psi$. The case when solutions provide us with only different $E$ in the set is called a non-degenerate case of solution, the degenerate case occurs when some of the $E$ in the set are equal, though these $E$ are corresponding to different eigenfunctions $\psi$.

## 4. A Simplification of the Solution of Equation (7) Done by Schrödinger

His simplified solution was usually based on a separation of the Hamiltonian $\hat{H}$ into two parts, namely

$$\hat{H} = \hat{H}_0 + \hat{H}^{\text{per}} \tag{8}$$

where the eigenequation

$$\hat{H}_0 \psi_0 = E_0 \psi_0 \tag{9}$$

is expected to be more simple to solve than that given in (7). Briefly a more simple Equation (9) is called the unperturbed equation with eigenvalues $E_0$ equal to the unperturbed energies and $\psi_0$ are called the unperturbed eigenfunctions. Because of (8) the perturbation potential entering the unperturbed Equation (9) is equal to:

$$\hat{H} - \hat{H}_0 = \hat{H}^{\text{per}} = V^{\text{per}}. \tag{10}$$

Usually it is assumed that

$$\hat{H}^{\text{per}} = V^{\text{per}}\left(\vec{r}\right), \tag{10a}$$

so (10a) is taken—for the sake of convenience—as equal to a term independent of the momentum operator, or operators. Having solutions of (9) we can calculate the matrix elements

$$\left\langle m \middle| V^{\text{per}} \middle| p \right\rangle = \int \psi_0^{(m)} V^{\text{per}} \psi_0^{(p)} \, \mathrm{d}V. \tag{11}$$

The matrix elements (11) combined with the eigenvalues $E_0$ entering (9) can provide us—according to the Schrödinger perturbation formalism—with the approximate energy eigenvalues of the more complicated eigenproblem (7).

This calculation can be done gradually for different perturbation orders $N$, beginning successfully with the lowest order $N = 1$. Huby [4] and Tong [5] calculated the number $S_N$ of kinds of the perturbation terms which should be built up from the matrix elements (11) for a given order $N$, on condition the perturbation concerns a non-degenerate quantum state. This number is equal to

$$S_N = \frac{(2N-2)!}{N!(N-1)!}. \tag{12}$$

But the derivation of a detailed shape of terms entering the number $S_N$ can be a complicated task, especially for large $N$. One of the aims of the present publication is to demonstrate that the perturbation calculation proposed by Schrödinger, especially its part referred to the $S_N$ terms, can be drastically simplified if the collision events of an originally unperturbed system with the per-

turbation potential

$$V^{\text{per}} = V^{\text{per}}\left(\vec{r}\right) \tag{13}$$

are arranged along a special scale of time. The scale has a circular-like shape and the number of collision kinds with the perturbation potential for each $N$ occurs precisely equal to $S_N$. Moreover, any collision kind, or collisions ensemble, is represented by a specified diagram created on the time scale. Simultaneously, the shape of the diagram provides us with a rule for calculating the corresponding contribution to the perturbation energy.

A final result for the perturbation energy obtained in this way for a given $N$ agrees with a corresponding energy obtained by the Schrödinger method. The details of calculations concerning the time scale and its applications are presented in the original author's papers; see [6]-[22].

## 5. Use of a Circular Scale of Time in the Schrödinger's Perturbation Problem

In fact only the perturbation of a non-degenerate Schrödinger quantum state $n$ was thoroughly considered with the aid of the mentioned scale. From the beginning of its application the circular scale of time was developed systematically for subsequent perturbation orders $N$:

$$N = 1, 2, 3, 4, \cdots \tag{14}$$

A physical meaning of $N$ was to give a number of collisions of an unperturbed system with the perturbation potential. This potential was usually assumed to depend solely on the position coordinate $\vec{r}$ of the particle:

$$\hat{H}^{\text{per}} = \hat{H}^{\text{per}}\left(\vec{r}\right) = V^{\text{per}}\left(\vec{r}\right) \tag{15}$$

Therefore the considered perturbation is independent of the time parameter $t$.

The number $N = 1$ refers to a single collision of the system with the perturbation (10), the number $N = 2$ refers to two collisions with $V^{\text{per}}\left(\vec{r}\right)$, etc. Any scale labelled by $N$ is assumed to be composed of the beginning-end (b.e.) point of time, in effect the scale represented by $N = 1$ has solely a single time point (b.e.) necessary for consideration.

The scale of $N = 2$ —giving the perturbation order 2—has two points of importance: beyond of a single beginning-end point it has the second point which refers to any non-perturbed state $p$ different than the considered unperturbed state $n$:

$$p \neq n. \tag{16}$$

The energy correction of state $n$ due to the perturbation of order $N = 1$ is represented by a single term

$$\Delta E_1 = \left\langle n \middle| V^{\text{per}} \middle| n \right\rangle. \tag{17}$$

On the other hand, the perturbation energy belonging to the order $N = 2$ is given by a sum

$$\Delta E_2 = \sum_p \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|n\rangle}{E_n^{(0)} - E_p^{(0)}} \tag{18}$$

where evidently the relation (16) does hold.

According to the formulae given by Tong and Huby [4] [5], the number of kinds of the perturbation terms entering order $N$ [see (12)] becomes:

$$S_1 = S_2 = 1 \tag{19}$$

which are in agreement with the number of $S_N$ given by the formula (12), see Figure 1 and Figure 2.

But in general we have

$$S_N > 1 \tag{20}$$

and our dominant interest is to calculate these $S_N$ terms.

In the case of $N = 1$ the diagram has only a single point—the beginning-end point—presented by Figure 1; for $N = 2$ the diagram is represented by Figure 2 having two isolated points. In the next step let us consider $N = 3$. In this case

$$S_3 = 2. \tag{21}$$

The circular scale for $N = 3$ has three points on it; see Figure 3. One point is the beginning-end (b.e.) point, the other time points are labelled on Figure 3 by the numbers 1 and 2.

The first of the $S_3$ terms in (9) is represented by the formula ( $p, q \neq n$ )

$$\Delta E_3^{(\text{part 1})} = \sum_p \sum_q \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|q\rangle\langle q|V^{\text{per}}|n\rangle}{\left(E_n^{(0)} - E_p^{(0)}\right)\left(E_n^{(0)} - E_q^{(0)}\right)}. \tag{22}$$

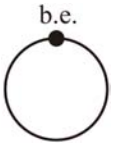



Figure 1. The diagram is corresponding to the perturbation order $N = 1$.


Figure 2. The diagram is corresponding to the perturbation order $N = 2$.

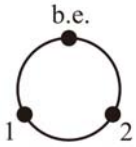



Figure 3. The diagram is corresponding to the perturbation order $N = 3$, part 1.

The second energy term dictated by (21) originates from contraction of the time points 1 and 2 which are present in Figure 3. This contraction 1:2 gives a diagram presented in Figure 4.

In fact the diagram in Figure 4 can be considered as representing the product of two terms: the first term is a time loop identical with the diagram characteristic for $N=1$ (see Figure 1), the other term is similar to the diagram for $N=2$ (see Figure 2). A difference from the term given in (22) is represented by the formula

$$\Delta E_3^{(\text{part 2})} = -\sum_p \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|n\rangle}{\left(E_n^{(0)} - E_p^{(0)}\right)^2}\Delta E_1. \tag{23}$$

In effect the full perturbation energy term for $N=3$ is

$$\Delta E^{(3)} = \Delta E_3^{(\text{part 1})} + \Delta E_3^{(\text{part 2})}. \tag{24}$$

The problem of sign attributed to $\Delta E_3^{(\text{part 1})}$ and $\Delta E_3^{(\text{part 2})}$ will be discussed below; see Section 6.

## 6. Abbreviated Formulae Applied in Calculating the Energy Perturbation Terms

The abbreviated formulae for $\Delta E_1$, $\Delta E_2$, $\Delta E_3^{(\text{first part})}$ and $\Delta E_3^{(\text{second part})}$ can be expressed as follows:

$$\Delta E_1 = \langle V\rangle, \tag{25}$$

$$\Delta E_2 = \langle VPV\rangle, \tag{26}$$

$$\Delta E_3^{(\text{first part})} = \langle VPVPV\rangle, \tag{27}$$

$$\Delta E_3^{(\text{second part})} = -\langle V\rangle\langle VP^2V\rangle, \tag{28}$$

where $\Delta E_1$ is given by (17), $\Delta E_2$ —by (18), $\Delta E_3^{(\text{part 1})}$ —by (22) and $\Delta E_3^{(\text{part 2})}$ —by (23).

A characteristic feature is that the case of $N=1$ has no $P$ terms entering $\Delta E_1$. The symbol $P$ in (26) refers to the ratio

$$\frac{1}{E_n^{(0)} - E_p^{(0)}} \tag{29}$$

entering only once for any state $p$ considered in the summation process in $\Delta E_2$. On the other hand $\Delta E_3$ has two kinds of $P$ terms, *viz.*

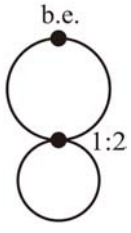



**Figure 4.** The diagram is corresponding to the perturbation order $N=3$, part 2.

$$\frac{1}{E_n^{(0)}-E_p^{(0)}},\frac{1}{E_n^{(0)}-E_q^{(0)}} \tag{30}$$

and product $P^2$ in (28) represents the term

$$\frac{1}{\left(E_n^{(0)}-E_p^{(0)}\right)^2} \tag{31}$$

entering the expression $\Delta E_3^{(\text{part 2})}$ in (23).

In general, a sum of powers $P$ entering any energy term $\Delta E_N$ should be equal to $N-1$.

The sign of $\Delta E_N$ is dependent on the number of terms entering the product representing a given $\Delta E_N$: an odd number of terms in the product implies a positive sign before it, so it is in the case of $\Delta E_1$, $\Delta E_2$ and $\Delta E_3^{(\text{part 1})}$; an even number of terms entering the product implies a negative sign [see $\Delta E_3^{(\text{part 2})}$ in (23)].

## 7. Contractions of the Time Points on the Scale of Time and the Number $S_N$

The number represented by $S_N$—and the formulae for the $S_N$ terms—can be obtained by considering the allowed contractions of the time points on the scale. An example of such contractions is given by the time points 1 and 2 represented by the symbol

$$1{:}2 \tag{32}$$

entering the time scale for $N=3$. No other contractions than (32) can be admitted for $N=3$ and its time scale.

But let us consider the time scale for $N=5$. In this case—beyond of the beginning-end time point (b.e.)—we have the time points

$$1,\ 2,\ 3,\ \text{and}\ 4 \tag{33}$$

on the scale; see Figure 5.

Since the point b.e. is excluded from contractions with the other time points, the allowed contractions to which the time points in Figure 5 can be submitted are:

$$\begin{aligned} &1{:}2,\quad 1{:}3,\quad 1{:}4,\quad 1{:}2{:}3,\quad 1{:}2{:}4,\quad 1{:}3{:}4,\\ &2{:}3,\quad 2{:}4,\quad 2{:}3{:}4,\\ &3{:}4,\quad 1{:}2{:}3{:}4. \end{aligned} \tag{34}$$

We see that the time points entering any contraction (34) should increase from smaller ones to larger ones.

But this property does not complete the limits of contractions choice. There are allowed also two combined contractions:

$$1{:}2 \text{ and } 3{:}4, \tag{35}$$

$$1{:}4 \text{ and } 2{:}3 \tag{36}$$

but there are n o t allowed contractions like

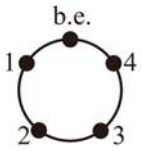



**Figure 5.** The time point b.e. and the free time points 1, 2, 3, and 4 on the scale.

$$1:3 \text{ and } 2:4. \tag{37}$$

The geometrical property which has to be satisfied for any set of contractions is that the lines (loops) associated with them should n o t cross.

In effect, together with the case of a single set of points given in (33), which are free from any contraction on the scale (see **Figure 5**), we have

$$1+6+5+2=14=S_N=S_5 \tag{38}$$

diagrams concerning points 1, 2, 3, and 4 when the points are submitted to contractions. Expression (38) is equal precisely to the result of the formula (12):

$$S_5=14. \tag{39}$$

The above procedure can be extended to an arbitrary perturbation order $N$. The order $N=7$ having $S_7=132$ terms was examined in [21] [22].

The diagrams representing the perturbation energies corresponding to contractions (34) are as in **Figure 6**.

The diagrams giving contractions (35) and (36) are as in **Figure 7**.

The diagram presented in **Figure 5** (having no contractions of the time points) gives the perturbation energy

$$\begin{aligned}\Delta E_5^{(1)} &= \langle VPVPVPVPV\rangle \\ &= \sum_p\sum_q\sum_r\sum_s \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|q\rangle\langle q|V^{\text{per}}|r\rangle\langle r|V^{\text{per}}|s\rangle\langle s|V^{\text{per}}|n\rangle}{\left(E_n^{(0)}-E_p^{(0)}\right)\left(E_n^{(0)}-E_q^{(0)}\right)\left(E_n^{(0)}-E_r^{(0)}\right)\left(E_n^{(0)}-E_s^{(0)}\right)}\end{aligned} \tag{40}$$

where

$$p,q,r,s\neq n. \tag{40a}$$

The energy terms represented by **Figure 6** are:

$$\begin{aligned}1:2 \text{ or } (\mathrm{a}) \to \Delta E_5^{(2)} &= -\langle VP^2VPVPV\rangle\langle V\rangle \\ &= -\sum_p\sum_q\sum_r \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|q\rangle\langle q|V^{\text{per}}|r\rangle\langle r|V^{\text{per}}|n\rangle}{\left(E_n^{(0)}-E_p^{(0)}\right)^2\left(E_n^{(0)}-E_q^{(0)}\right)\left(E_n^{(0)}-E_r^{(0)}\right)}\Delta E_1\end{aligned} \tag{41}$$

$$\begin{aligned}2:3 \text{ or } (\mathrm{g}) \to \Delta E_5^{(3)} &= -\langle VPVP^2VPV\rangle\langle V\rangle \\ &= -\sum_p\sum_q\sum_r \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|q\rangle\langle q|V^{\text{per}}|r\rangle\langle r|V^{\text{per}}|n\rangle}{\left(E_n^{(0)}-E_p^{(0)}\right)\left(E_n^{(0)}-E_q^{(0)}\right)^2\left(E_n^{(0)}-E_r^{(0)}\right)}\Delta E_1\end{aligned} \tag{42}$$

$$\begin{aligned}3:4 \text{ or } (\mathrm{j}) \to \Delta E_5^{(4)} &= -\langle VPVPVP^2V\rangle\langle V\rangle \\ &= -\sum_p\sum_q\sum_r \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|q\rangle\langle q|V^{\text{per}}|r\rangle\langle r|V^{\text{per}}|n\rangle}{\left(E_n^{(0)}-E_p^{(0)}\right)\left(E_n^{(0)}-E_q^{(0)}\right)\left(E_n^{(0)}-E_r^{(0)}\right)^2}\Delta E_1\end{aligned} \tag{43}$$

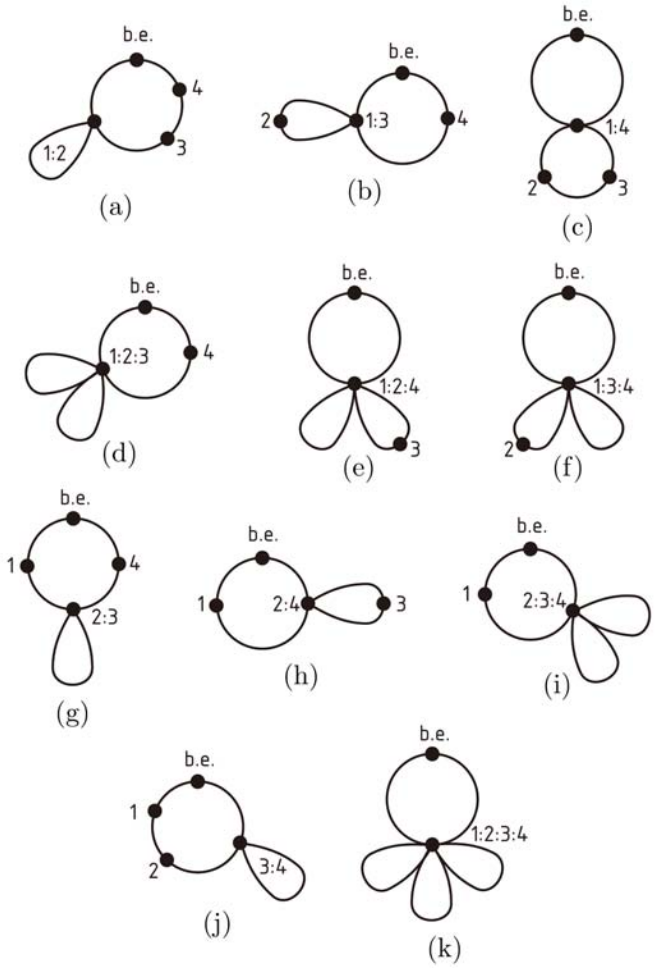



**Figure 6.** The diagrams representing the b.e. time points and contractions (34) on the time scale.

$$1:3 \text{ or } (\text{b}) \to \Delta E_5^{(5)} = -\left\langle VP^2VPV \right\rangle \left\langle VPV \right\rangle$$
$$= -\sum_p \sum_q \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|q\rangle\langle q|V^{\text{per}}|n\rangle}{\left(E_n^{(0)} - E_p^{(0)}\right)^2 \left(E_n^{(0)} - E_q^{(0)}\right)} \Delta E_2 \tag{44}$$

$$1:4 \text{ or } (\text{c}) \to \Delta E_5^{(6)} = -\left\langle VP^2V \right\rangle \left\langle VPVPV \right\rangle$$
$$= -\sum_p \frac{\langle n|V^{\text{per}}|p\rangle\langle p|V^{\text{per}}|n\rangle}{\left(E_n^{(0)} - E_p^{(0)}\right)^2} \Delta E_3^{(\text{part 1})}. \tag{45}$$

In the next step

$$1:4 \cap 2:3 \to \Delta E_5^{(6')} = \left\langle VP^2V \right\rangle \left\langle VP^2V \right\rangle \left\langle V \right\rangle \tag{45a}$$

[see diagram (b') in **Figure 7**] gives together with $\Delta E_5^{(6)}$ the result:

$$\Delta E_5^{(6)} + \Delta E_5^{(6')} = -\left\langle VP^2V \right\rangle \left[ \left\langle VPVPV \right\rangle - \left\langle VP^2V \right\rangle \left\langle V \right\rangle \right]$$
$$= -\left\langle VP^2V \right\rangle \left[ \Delta E_3^{(\text{part 1})} + \Delta E_3^{(\text{part 2})} \right] \tag{46}$$
$$= -\left\langle VP^2V \right\rangle \Delta E_3;$$

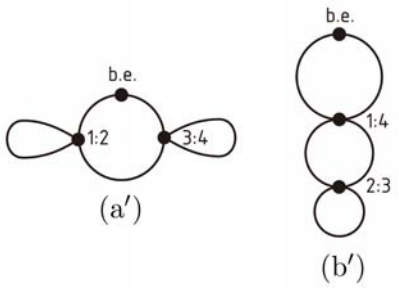



**Figure 7.** The diagrams representing the b.e. time points and contractions (35) and (36) on the time scale.

see (22) and (23). This implies that (46) provides us with two Schrödinger perturbation terms for energy.

The remaining contractions of the time points give:

$$1:2:3 \text{ or } (\mathrm{d}) \to \Delta E_5^{(7)} = \left\langle VP^3VPV \right\rangle \left( \left\langle V \right\rangle \right)^2$$
$$= \sum_p \sum_q \frac{\left\langle n \left| V^{\mathrm{per}} \right| p \right\rangle \left\langle p \left| V^{\mathrm{per}} \right| q \right\rangle \left\langle q \left| V^{\mathrm{per}} \right| n \right\rangle}{\left( E_n^{(0)} - E_p^{(0)} \right)^3 \left( E_n^{(0)} - E_q^{(0)} \right)} \left( \Delta E_1 \right)^2 \tag{47}$$

$$1:2:4 \text{ or } (\mathrm{e}) \to \Delta E_5^{(8)} = \left\langle VP^3VPV \right\rangle \left\langle V \right\rangle \left\langle VPV \right\rangle$$
$$= \sum_p \frac{\left\langle n \left| V^{\mathrm{per}} \right| p \right\rangle \left\langle p \left| V^{\mathrm{per}} \right| n \right\rangle}{\left( E_n^{(0)} - E_p^{(0)} \right)^3} \Delta E_1 \Delta E_2 \tag{48}$$

$$1:3:4 \text{ or } (\mathrm{f}) \to \Delta E_5^{(9)} = \left\langle VP^3VPV \right\rangle \left\langle VPV \right\rangle \left\langle V \right\rangle$$
$$= \sum_p \frac{\left\langle n \left| V^{\mathrm{per}} \right| p \right\rangle \left\langle p \left| V^{\mathrm{per}} \right| n \right\rangle}{\left( E_n^{(0)} - E_p^{(0)} \right)^3} \Delta E_2 \Delta E_1 \tag{49}$$

$$2:3:4 \text{ or } (\mathrm{i}) \to \Delta E_5^{(11)} = \left\langle VPVP^3V \right\rangle \left( \left\langle V \right\rangle \right)^2$$
$$= \sum_p \sum_q \frac{\left\langle n \left| V^{\mathrm{per}} \right| p \right\rangle \left\langle p \left| V^{\mathrm{per}} \right| q \right\rangle \left\langle q \left| V^{\mathrm{per}} \right| n \right\rangle}{\left( E_n^{(0)} - E_p^{(0)} \right) \left( E_n^{(0)} - E_q^{(0)} \right)^3} \left( \Delta E_1 \right)^2 \tag{50}$$

$$1:2:3:4 \text{ or } (\mathrm{k}) \to \Delta E_5^{(12)} = -\left\langle VP^3VPV \right\rangle \left( \left\langle V \right\rangle \right)^2$$
$$= -\sum_p \frac{\left\langle n \left| V^{\mathrm{per}} \right| p \right\rangle \left\langle p \left| V^{\mathrm{per}} \right| n \right\rangle}{\left( E_n^{(0)} - E_p^{(0)} \right)^4} \left( \Delta E_1 \right)^3 . \tag{51}$$

The energy term represented by (a') in Figure 7 is

$$\Delta E_5^{(13)} = \left\langle VP^2VP^2V \right\rangle \left( \left\langle V \right\rangle \right)^2$$
$$= \sum_p \sum_q \frac{\left\langle n \left| V^{\mathrm{per}} \right| p \right\rangle \left\langle p \left| V^{\mathrm{per}} \right| q \right\rangle \left\langle q \left| V^{\mathrm{per}} \right| n \right\rangle}{\left( E_n^{(0)} - E_p^{(0)} \right)^2 \left( E_n^{(0)} - E_q^{(0)} \right)^2} \left( \Delta E_1 \right)^2 . \tag{52}$$

Equation (40a) should be satisfied in all summations.

## 8. Summary of Results

The main idea of the paper presented by the author—and in former his papers

quoted here—was to demonstrate that a rather tedious approach to the Schrödinger perturbation energy can be much simplified when a circular scale of time in classifying the perturbations events is applied. As a special case a perturbation of a non-degenerate quantum state is examined.

First the number of terms entering the calculation is in a perfect agreement with the number of the formulae expected by Huby and Tong. The next point is that an arrangement of the time points on the scale gives a ready mathematical access to the formulae entering the Schrödinger perturbation theory for a given perturbation order $N$. This is a convenient situation because—for example in applying the Feynmann diagrams—the number of the perturbation terms and their character are much larger and more complicated than those necessary to perform the proper Schrödinger calculations for a chosen $N \gg 1$.

In references [6]-[22] are given the diagrams of the time scale and their applications corresponding to the Schrödinger perturbation terms whose orders do not exceed $N = 7$. Some philosophical repercussions concerning the shape of the applied time scale and its use are also presented. In [9] a rather thorough comparison is done between the Feynman approach based on an infinite (straight-linear) time scale and a circular-time approach to the Schrödinger perturbation theory. Another comparison done in [22] does refer to the present perturbation calculation to the Leibniz theory and that outlined in the Ziman book [23].

The arrangement of the perturbation events on a special, *viz.* circular, scale of time, allowed us to obtain the perturbation energies of a non-degenerate quantum system without solving the corresponding perturbation equations.

## Acknowledgements

I am extremely grateful to Ms. Agnieszka Swiatkiewicz for her kind help in preparing the texts of my manuscripts.

## Conflicts of Interest

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

## References

[1] Sommerfeld, A. (1949) Mechanik. 4th Edition, Akademische Verlagsgesellschaft, Leipzig.

[2] Landau, L.D. and Lifshits, E.M. (1948) Field Theory. 2nd Edition, OGIZ, Moscow. (In Russian)

[3] Sommerfeld, A. (1939) Atombau und Spektrallinien. 5th Edition, Vieweg, Braunschweig.

[4] Huby, R. (1961) *Proceedings of the Physical Society* (*London*), **78**, 529-536. https://doi.org/10.1088/0370-1328/78/4/306

[5] Tong, B.Y. (1962) *Proceedings of the Physical Society* (*London*), **80**, 1101-1104. https://doi.org/10.1088/0370-1328/80/5/308

[6] Olszewski, S. (1991) *Zeitschrift fuer Naturforschung A*, **46**, 313-320. https://doi.org/10.1515/zna-1991-0404

[7] Olszewski, S. (1992) *Studia Philosophiae Christianae*, **28**, 119-135.

[8] Olszewski, S. (1994) *Studia Philosophiae Christianae*, **30**, 203-207.

[9] Olszewski, S. and Kwiatkowski, T. (1998) *Computers in Chemistry*, **22**, 445-464. https://doi.org/10.1016/S0097-8485(98)00023-0

[10] Olszewski, S. (2004) *International Journal of Quantum Chemistry*, **97**, 784-801. https://doi.org/10.1002/qua.10776

[11] Olszewski, S. (2003) *Trends in Physical Chemistry*, **9**, 69-101.

[12] Olszewski, S. (2011) *Journal of Quantum Information Science*, **1**, 142-148. https://doi.org/10.4236/jqis.2011.13020

[13] Olszewski, S. (2013) *Quantum Matter*, **2**, 481-483. https://doi.org/10.1166/qm.2013.1085

[14] Olszewski, S. (2015) *Reviews in Theoretical Science*, **3**, 1-15. https://doi.org/10.1166/rits.2015.1029

[15] Olszewski, S. (2015) *Quantum Matter*, **4**, 523-532. https://doi.org/10.1166/qm.2015.1227

[16] Olszewski, S. (2014) *Journal of Modern Physics*, **5**, 1502-1523. https://doi.org/10.4236/jmp.2014.515152

[17] Olszewski, S. (2014) *Journal of Quantum Information Science*, **4**, 269-283. https://doi.org/10.4236/jqis.2014.44022

[18] Olszewski, S. (2017) *Journal of Modern Physics*, **8**, 1650-1685. https://doi.org/10.4236/jmp.2017.89098

[19] Olszewski, S. (2018) *Journal of Modern Physics*, **9**, 1491-1521. https://doi.org/10.4236/jmp.2018.98093

[20] Olszewski, S. (2019) *Advances in Pure Mathematics*, **9**, 228-241. https://doi.org/10.4236/apm.2019.93011

[21] Olszewski, S. (2019) *World Journal of Mechanics*, **9**, 113-145. https://doi.org/10.4236/wjm.2019.95009

[22] Olszewski, S. (2020) *Journal of Modern Physics*, **11**, 1536-1558. https://doi.org/10.4236/jmp.2020.1110095

[23] Ziman, J.-M. (1969) Elements of Advanced Quantum Theory. Cambridge University Press, London.

Journal of Modern Physics, 2022, 13, 1093-1098
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Antiproton Production with a Fixed Target and Search for Superheavy Particles at the LHC

**Alexey B. Kurepin[1*], Nikolay A. Kurepin[2], Konstantin A. Skazytkin[2]**

[1]Institute for Nuclear Research, Moscow, Russia
[2]Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
Email: *kurepin@inr.ru

**How to cite this paper:** Kurepin, A.B., Kurepin, N.A. and Skazytkin, K.A. (2022) Antiproton Production with a Fixed Target and Search for Superheavy Particles at the LHC. *Journal of Modern Physics*, **13**, 1093-1098.
https://doi.org/10.4236/jmp.2022.137062

**Received:** May 9, 2022
**Accepted:** July 15, 2022
**Published:** July 18, 2022



## Abstract

A proposal for an experiment to measure the cross section of antiproton production in a proton-nuclear collision in a kinematically forbidden region for nucleon-nucleon interaction on a fixed LHC target is considered. It is shown that this process can be separated from the kinematically allowed production process using the existing detectors of the ALICE facility at a proton energy of 7 TeV with a fixed nuclear target. Assuming the scale dependence of the cross section, the data obtained can be used to estimate the subthreshold cross section for the production of superheavy particles with a mass of several tens of TeV in the LHC lead nucleus beam.

## Keywords

Antiproton, Fixed Target, Collider, Heavy Ion Collision, Superheavy Particles

## 1. Introduction

Among the processes of particle production in nuclear collisions at high energy beams, one can highlight those where production occurs under kinematic conditions that are forbidden in the nucleon-nucleon interaction. At present, this problem is of great importance for studying the possibility of particle production with beams of lead nuclei for a mass exceeding 14 TeV, which is attainable in a nucleon-nucleon collision at the Large Hadron Collider (LHC). Actually, the energy at the center of mass in the collision of lead nuclei at the LHC at a beam energy of 2.76 TeV per nucleon is 1150 TeV. The existence of such superheavy particles with a mass much less than the Planck mass is forbidden in the Grand Unification Theory. However, when trying to solve the problem of the hierarchy of interactions in models of superstrings, super symmetry and introducing additional dimensions, the creation of such particles is allowed. For example, in the

model of extra dimensions with a compactification radius of one Fermi, particles with a mass of several tens of TeV can be produced [1] [2].

Obviously, for the production of superheavy particles, the cross section of the "subthreshold" process will be very small. Theoretical estimates of this cross section are apparently impossible at present. However, an analogy can be drawn with the subthreshold production of antiprotons at intermediate energies, since the subthreshold process is obviously associated with the correlation of nucleons or quarks in colliding nuclei and could be weakly dependent on energy. Nevertheless, such a study must also be carried out at higher energies.

This investigation can be carried out in the study of the production process under kinematic conditions that are forbidden in the nucleon-nucleon interaction. However, obtaining the necessary data in the collider operation mode is impossible. In this case, the produced antiprotons have an energy of several hundred GeV, and the spectrometry of antiprotons at such energies with existing detectors does not provide the required accuracy. As first shown in this article, it is possible to measure the production of antiprotons in a proton beam with an energy of 7 TeV at the LHC on a fixed target of heavy nuclei ( $\sqrt{s} = 114.7$ GeV ). Such a study was started at the U-70 accelerator in Protvino [3] at an energy of 19 GeV ( $\sqrt{s} = 6.05$ GeV ) per nucleon on carbon nuclei. However, the measurement of antiproton production was carried out at small forward angles, which leads to large antiproton momenta and less accuracy. The measurement of the transverse momentum of antiprotons proposed in this article makes it possible to obtain better accuracy at a much higher energy.

## 2. Subthreshold Production of Antiprotons in Nuclear Collisions

The study of the production of antiprotons in proton-nuclear and nucleus-nucleus collisions at energies below the production threshold in a nucleon-nucleon collision has been the subject of a significant number of works [4] [5] [6] [7] [8]. The measurements were carried out with proton and nuclear beams at JINR, BNL, GSI and KEK. In all experiments, the values of the production cross sections were obtained, which significantly exceeded the estimates obtained when taking into account the lowering of the production threshold due to the Fermi motion of nucleons in the nucleus [9]. A unified phenomenological description of all experimental data was obtained in the generalized parton model [10] [11]. The model is based on the introduction of the parton distribution parameter not only in the target nucleus ($x$), but also in the incident nucleus ($z$). Due to the conservation of the 4-momentum, we obtain the following results:

$$\left(zP_1 + xP_2 - P_d\right)^2 = \left(zP_1' + xP_2' + P_i\right)^2 \tag{1}$$

where $P_1$, $P_2$ and $P_d$—are the 4-momenta of the incident and fixed nucleons in the nuclei and of the produced antiproton respectively, and $P_i$ is the 4-momentum of an additional proton for conservation of the baryon number. For the maximum values of the momentum of the produced antiproton:

$$\overrightarrow{p_1'} = \overrightarrow{p_2'} = \overrightarrow{p_i'} = 0 \tag{2}$$

We have: $\left(P_1'\right)^2 = m^2$ and $\left(P_2'\right)^2 = m^2$, where $m$—is the mass of a nucleon, $\overrightarrow{p_1'}$, $\overrightarrow{p_2'}$, $\overrightarrow{p_i'}$ momenta of incident and fixed nucleons in nuclei and an additional proton.

From Equation (1), the mutual dependence of the parameters $x$ and $z$ can be obtained [10]. The production of antiprotons in nucleus-nucleus collisions is now possible not only at small values of the parton parameters, but also in the kinematically forbidden region for nucleon-nucleon interactions at $x > 1$ and $z > 1$.

In reference [11], a universal dependence of the scaling type of all subthreshold data on the production of antiprotons was obtained at $x$ in the range of 1 - 4 and at $z$ values equal to 1 for incident protons, 1.3 for deuterons, 2 for carbon nuclei, and 3 for more heavy nuclei at energies from 2 to 6 GeV per nucleon:

$$\left(A_1 A_2\right)^{-0.43} \cdot E_1 \frac{\mathrm{d}^3\sigma}{\mathrm{d}p^3}\left[\mathrm{mb}\cdot\mathrm{GeV}^{-2}\cdot\mathrm{c}^3\cdot\mathrm{sr}^{-1}\right] = 0.57\exp\left(-x/0.158\right) \tag{3}$$

where $E_1$ is the total energy of antiproton and $A_1$ and $A_2$ are the mass numbers of colliding nuclei.

This dependence has an exponential form and describes the data for proton-nuclear collisions well. For nucleus-nuclear collisions, the deviations of individual data from the curve are larger, but the range of cross sections reaches four orders of magnitude.

Assuming a weak dependence of the obtained dependence of the subthreshold production cross section on energy, the yield of superheavy particles with a mass of 16 TeV on the LHC proton beam in the collision of lead nuclei was estimated [12]. The obtained value of approximately 70 particles per year allows planning the corresponding experiment. However, for more reasonable estimates, it is necessary to determine the dependence of the production cross section on the scaling parameters in the subthreshold process, at values $x > 1$, at high energies closer to the LHC energies.

In the next section, we analyse the possibility of measuring the cross section for antiproton production in the kinematically forbidden region on a fixed target of the LHC collider.

## 3. Production of Antiprotons in a Kinematically Forbidden Region at a Fixed Target of the LHC Collider

To determine the possibility of measuring the yield of antiprotons in the kinematically forbidden region on a fixed target of the LHC collider [13] the kinematics of antiproton production at an energy of 7 TeV on bismuth nuclei were calculated. In **Figure 1**, the magnitudes of the maximum transverse momentum values of antiprotons in the center-of-mass system are given for the parameters $x = 1$ and $x = 2$, depending on the pseudorapidity. These magnitudes differ significantly. Thus, it is possible to separate the kinematically allowed process from the forbidden process in the nucleon-nucleon collision.

The identification and measurement of the transverse momentum in the ALICE installation is carried out by the TPC projection camera. The fixed target will be located at a distance of 480 cm from the IP. The geometry of the TPC and the beam tube limits the range of possible angles from 5 to 28 degrees. Figure 2 and Figure 3 show the corresponding intervals of pseudorapidity for $x = 1$ and $x = 2$. The required transverse momentum ranges of 3 - 20 GeV are available for TPC measurements [14].

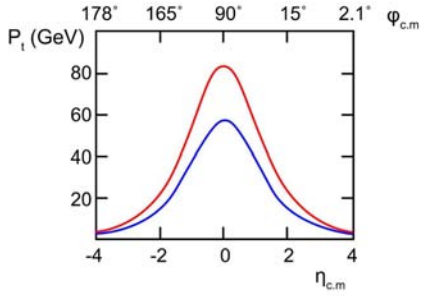



Figure 1. Dependence of the transverse momentum of antiprotons on the pseudorapidity in the center of mass. Blue line with $x = 1$, red line with $x = 2$.

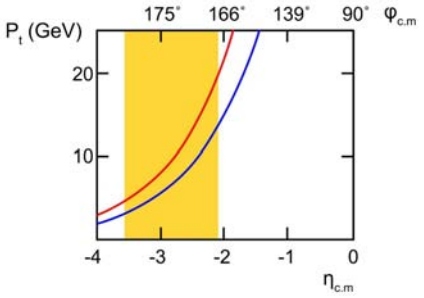



Figure 2. Dependence of the transverse momentum of antiprotons on the pseudorapidity in the center of mass. Blue line with $x = 1$, red line with $x = 2$. The area available with a fixed target and $x = 1$ is marked in yellow.

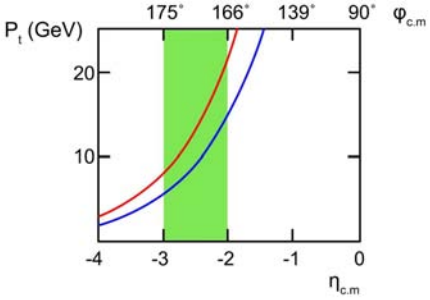



Figure 3. Dependence of the transverse momentum of antiprotons on the pseudorapidity in the center of mass. Blue line with $x = 1$, red line with $x = 2$. The area available with a fixed target and $x = 2$ is marked in green.

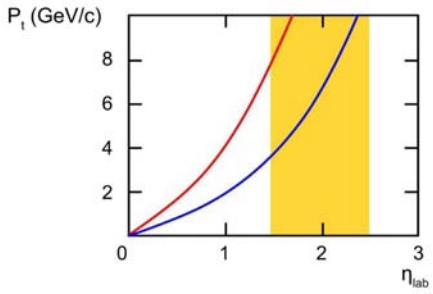



**Figure 4.** Dependence of the transverse momentum of antiprotons on the pseudorapidity in a laboratory system. Blue line with $x = 1$, red line with $x = 2$. The area available with a fixed target is marked in yellow.

**Table 1.** Parameter $x$, antiproton production cross sections and antiproton yield as a function of the antiproton transverse momentum.

| $P_t$ | 4 | 6 | 8 | GeV |
|---|---|---|---|---|
| $E_d$ | 8.5 | 12.8 | 17 | GeV |
| $x$ | 1.1 | 1.6 | 2.18 | - |
| $\sigma_{inv}$ | $8 \times 10^{-3}$ | $6 \times 10^{-4}$ | $8 \times 10^{-6}$ | mb·GeV$^{-2}$·c$^3$·sr$^{-1}$ |
| $N_d$ | $25 \times 10^3$ | $3 \times 10^3$ | 50 | 1/hour |

**Figure 4** shows the dependence of the transverse momentum on the pseudorapidity in the laboratory system.

When analysing the measured cross sections for the production of antiprotons, the parameters $x$ and $z$ must satisfy the following relation:

$$x = \frac{z \cdot E \cdot E_d \cdot \left(1 - \cos(\theta)\right) - M^2}{\left(z \cdot E - E_d\right) m} \tag{4}$$

where $E$ is the beam energy, $E_d$ is the antiproton energy, $M$ is the antiproton mass, $m$ is the nucleon mass and $\theta$ is the production angle.

At a high energy of 7 TeV of a proton beam, the parameter $x$ is practically independent of the parameter $z$ and beam energy:

$$x \approx \frac{E_d \cdot \left(1 - \cos(\theta)\right)}{m} \tag{5}$$

For the parameters of the planned experiment: $\theta = 28^0$, $\Delta p = 1$ GeV, $\Delta\Omega = 0.1$ sr, the values of parameter $x$, as a function of the antiproton transverse momentum, are given in **Table 1**. The same table shows the production cross section calculated by formula (3) and antiproton yield at a luminosity of $10^{30}$ см$^{-2}$·сек$^{-1}$.

## 4. Conclusion

Investigation of antiproton production in the kinematically forbidden region in the nucleon-nucleon interaction on a fixed target of the LHC collider is possible with the existing ALICE facility detectors. The data obtained on the dependence

of the subthreshold production cross section on the scaling parameter $x > 1$ can be used to estimate the yield of superheavy particle production with the LHC lead nucleus beam. The results of the proposed measurements will allow studying the dependence of the scaling effect on energy, as well as obtaining new data on collective effects in nuclei.

## Acknowledgements

We thank the members of the fixed-target ALICE study group and the AFTER study group for helpful discussions and interest in the work.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Arkani-Hamed, N., Dimopoulos, S. and Dvali, G. (1998) *Physics Letters B*, **429**, 263. https://doi.org/10.1016/S0370-2693(98)00466-3

[2] Rubakov, V.A. (2003) *Physics-Uspekhi*, **46**, 211. https://doi.org/10.1070/PU2003v046n02ABEH001355

[3] Afonin, A.G., *et al.* (2020) *Physics of Atomic Nuclei*, **83**, 140. https://doi.org/10.31857/S0044002720020014

[4] Baldin, A.A., *et al.* (1988) *JETP Letters*, **48**, 137-140.

[5] Baldin, A.A., *et al.* (1990) *Nuclear Physics* A, **519**, 407-411. https://doi.org/10.1016/0375-9474(90)90644-2

[6] Carroll, J.B., *et al.* (1989) *Physical Review Letters*, **62**, 1829-1832. https://doi.org/10.1103/PhysRevLett.62.1829

[7] Chiba, J., *et al.* (1993) *Nuclear Physics* A, **553**, 771-774. https://doi.org/10.1016/0375-9474(93)90696-U

[8] Schroeter, A., *et al.* (1993) *Nuclear Physics* A, **553**, 775-778. https://doi.org/10.1016/0375-9474(93)90697-V

[9] Shor, A., Perez-Mendez, V. and Ganezer, K. (1990) *Nuclear Physics A*, **514**, 717-733. https://doi.org/10.1016/0375-9474(90)90019-I

[10] Stavinski, V.S. (1988) JINR Rapid Comm. № 18-86, p. 5.

[11] Kurepin, A.B., Shileev, K.A. and Topilskaya, N.S. (1997) *Genshiryoku Kenkyu, Tokyo*, **41**, 177-182.

[12] Kurepin, A. (2021) *Journal of Modern Physics*, **12**, 433-439. https://doi.org/10.4236/jmp.2021.124030

[13] Hadjidakis, C., *et al.* (2021) *Physics Reports*, **911**, 1-83. https://doi.org/10.1016/j.physrep.2021.01.002

[14] Acharya, S., *et al.*, ALICE Collaboration (2021) *Journal of High Energy Physics*, No. 5, Article 290. https://doi.org/10.1007/JHEP05(2021)290

Journal of Modern Physics, 2022, 13, 1099-1116
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# About the Observed Asymmetry between Matter and Antimatter

**Luis Grave de Peralta**[1,2]

[1]Department of Physics and Astronomy, Texas Tech University, Lubbock, TX, USA
[2]Nano Tech Center, Texas Tech University, Lubbock, TX, USA
Email: luis.grave-de-peralta@ttu.edu

**How to cite this paper:** Grave de Peralta, L. (2022) About the Observed Asymmetry between Matter and Antimatter. *Journal of Modern Physics*, **13**, 1099-1116.
https://doi.org/10.4236/jmp.2022.137063

**Received:** June 10, 2022
**Accepted:** July 15, 2022
**Published:** July 18, 2022



## Abstract

Some consequences, due to the existence of a pair of decoupled Schrödinger-like but relativistic quantum mechanics wave equations, are explored. It is shown that one equation directly describes the quantum states of a single spin-0 particle, and the other one indirectly describes the quantum states of the corresponding antiparticle. In correspondence with the matter-antimatter symmetry, for a Coulomb potential, a charge conjugation operation transforms the second equation in the first one. However, if a particle could interact with itself (gravitationally or electrically) due to the spread of its wavefunction, the C-symmetry could be broken; therefore, matter and antimatter could be distinguished. Under these assumptions, it is deducted the impossibility of the existence of particles and antiparticles with a mass larger than the Plank mass ($m_P$), or with the absolute value of the charge larger than the Plank charge ($q_P$). It is proposed the existence of primordial antimatter electrical sinks. It is also suggested that all macroscopic matter objects with a mass $m > m_P$, and all macroscopic antimatter bodies with a charge $|q| > q_P$ should not be quantum but classical objects. It is argued that these findings could explain the absence of antimatter with a complicated structure and partially explain the excess of charged matter in the known Universe.

## Keywords

Relativistic Quantum Mechanics, Antimatter, Primordial Black Holes

## 1. Introduction

We do not live in an antimatter world but in a matter one. The observed asymmetry between matter and antimatter has two aspects. First, matter seems to be much more abundant in Nature than antimatter. Second, we are not surrounded by antimatter bodies with a complicated structure but by complicated matter

objects like matter atoms, molecules, and bulky bodies formed by them. It has been speculated that the second is a consequence of the first, and the first happens due to a yet unknown cause at the beginning of the known Universe [1] [2]. The so-called charge-parity-time reversal (CPT) symmetry justifies the current faith in such explanation. This is because we believe all laws of Physics should be CPT invariable [1] [2] [3], and because it is a general belief that the CPT symmetry implies that a matter world should be undistinguishable from the antimatter one [1] [2] [3].

In this work, it is explored a different explanation of why we are not surrounded by antimatter bodies with a complicated structure but by complicated matter bodies, and it is partially addressed the observed abundance of matter in Nature. For simplicity, only electrostatic and (Newtonian) gravitational interactions are considered in this work. The author invites the reader to start our journey in the general Grave de Peralta (gGP) equation [4] [5]:

$$i\hbar\frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2\mu}\nabla^2\Psi + V\Psi. \tag{1}$$

This is a Schrödinger-like, but relativistic quantum mechanics equation, that describes the quantum states of a particle moving in a scalar potential ($V$) with an effective mass $\mu = \mu_+ > 0$, which depends on the parameter $\gamma$ in the following way [4]:

$$\mu_+ = \frac{1+\gamma}{2}m > 0. \tag{2}$$

When $\gamma = 1$, then $\mu_+ = m$, the relativistic invariant mass of the particle; thus, Equation (1) coincides with the Schrödinger equation [6] [7]. Equation (1) is the Grave de Peralta (GP) equation when $\gamma$ is the Lorentz factor of special theory of relativity [5] [8] [9] [10]. Equation (1) with $\mu = \mu_+ > 0$ has been successfully used for extending, to the relativistic domain, the non-relativistic results previously obtained using the Schrödinger equation [4] [5] [8]-[15]. Equation (1) was also used for demonstrating the impossibility of the existence of elemental quantum particles with a mass larger than the Plank mass ($m_P \approx 20\ \mu g$) [5]. In addition, it was also suggested that all matter bodies with $m > m_P$ should not be quantum but classical objects [5]. These results were obtained using the hypothesis that, due to the spread of its mass density through its wavefunction, a matter particle can gravitationally interact with itself [5] [16].

In this work, for the first time, the consequences of considering $\mu = \mu_- = -\mu_+ < 0$ are explored. First, it is shown that Equation (1), but with $\mu = \mu_-$, indirectly describes the quantum states of the antiparticle that is associated to a particle. The quantum (particle) states associated to this particle are described by Equation (1) with $\mu = \mu_+ > 0$. Second, Equation (1) is used, but with $\mu = \mu_- < 0$, in combination with the hypothesis that a particle can electrostatically interact with itself, due to the spread of its charge density through its wavefunction, for demonstrating the impossibility of the existence of elemental quantum antiparticles with the absolute value of its charge ($|q|$) larger than the Plank charge, $q_P \approx$

11$e$, where $e$ is the charge of a proton. It is also suggested that all antimatter bodies with $|q| > q_P$ should not be quantum but classical objects.

Finally, it is proposed the possible existence of primordial antimatter electrical sinks, which may have formed when primordial quantum fluctuations of charged antimatter, with $|q| > q_P$, occurred at the beginning of the times. The existence of such primordial antimatter electrical sinks may partially explain the observed excess of charged matter in the known Universe. Moreover, it is suggested that we do not observe antimatter bodies with a complicated structure because antimatter atoms and molecules only can be formed when antimatter nuclei and positron clouds are quantum objects. However, they may be classical when $|q| > q_P$.

## 2. The gGP Equations

Like the Klein-Gordon equation, Equation (1) can be obtained, after a formal first quantization procedure, from the Lorentz-invariance of the magnitude of the four-component energy-momentum vector corresponding to a classical particle of mass $m$, total energy $E$, and three-component linear momentum $\boldsymbol{p}$, which is moving in a scalar potential $V$ [3] [17] [18]:

$$\sqrt{\frac{(E-V)^2}{c^2}-\boldsymbol{p}^2}=mc. \tag{3}$$

In Equation (3), $c$ is the speed of the light in vacuum, and $(E-V)$ is the sum of the kinetic energy of the particle ($K$) plus the energy associated to its mass ($mc^2$). Solving Equation (3) for $(E-V)$, we obtain two possible values of $(E-V)$ corresponding to each value of $\boldsymbol{p}$:

$$E-V=\pm\sqrt{\boldsymbol{p}^2c^2+m^2c^4}=\pm\gamma mc^2,\ \text{with}\ \gamma=\sqrt{1+\frac{p^2}{m^2c^2}}. \tag{4}$$

It is worth noting that $\gamma=1$ means the particle is at rest. In what follows, we will add the subindexes (+) and (–) to all the magnitudes related with the positive and negative values of $(E-V)$, respectively. When $(E-V)=+\gamma mc^2>0$, the relation between $(E-V)$ and $K$ is:

$$E_+-V=+\gamma mc^2=K_++mc^2\Rightarrow K_+=(\gamma-1)mc^2\geq 0. \tag{5}$$

However, when $(E-V)=-\gamma mc^2<0$, the relation between $(E-V)$ and $K$ should be such that $K_-=0$ when $\gamma=1$; therefore:

$$E_--V=-\gamma mc^2=K_--mc^2\Rightarrow K_-=(1-\gamma)mc^2=-K_+\leq 0. \tag{6}$$

Defining $E'$ such that:

$$E'_\pm-V=\begin{matrix}(E_+-V)-mc^2=K_+\\(E_--V)+mc^2=K_-\end{matrix}\Rightarrow E'_\pm-V=\begin{matrix}K_+=(\gamma-1)mc^2\geq 0\\K_-=(1-\gamma)mc^2=-K_+\leq 0\end{matrix}. \tag{7}$$

And using Equation (4) for evaluating $\gamma^2$, we can obtain that $K_+$ and $K_-$ are also given by the following equations:

$$\begin{aligned}&(\gamma+1)K_{\pm}=\pm(\gamma+1)\left[(\gamma-1)mc^2\right]=\pm\left(\gamma^2-1\right)mc^2=\pm\frac{\boldsymbol{p}^2}{m}\\&\Rightarrow K_{\pm}=\pm\frac{\boldsymbol{p}^2}{(1+\gamma)m}\end{aligned}\tag{8}$$

Therefore, combining Equations (7) and (8), we obtain:

$$E'_{\pm}=\pm\frac{\boldsymbol{p}^2}{(1+\gamma)m}+V.\tag{9}$$

Now, by making in Equation (9) the following formal first-quantization substitutions:

$$E'\to i\hbar\frac{\partial}{\partial t},\ \ \boldsymbol{p}\to-i\hbar\boldsymbol{\nabla}.\tag{10}$$

We obtain a pair of Poirier-Grave de Peralta (PGP) equations [5]:

$$i\hbar\frac{\partial}{\partial t}\Psi_{\pm}=\mp\frac{\hbar^2}{(1+\hat{\gamma})m}\nabla^2\Psi_{\pm}+V\Psi_{\pm},\ \text{with}\ \hat{\gamma}=\sqrt{1+\frac{\hat{\boldsymbol{p}}^2}{m^2c^2}}.\tag{11}$$

Readers familiar with high energy physics may recognize that the PGP equation for $\Psi_+$ is related with a particular case of the spinless Salpeter equation [19] [20], which is a known Lorentz-covariant alternative for the Klein-Gordon equation. Like the spinless Salpeter equation, the PGP equations are Lorentz-covariant; thus, valid relativistic quantum mechanics equations [4] [5]. Different approaches may be used for removing the operator $\gamma$ in Equation (11) and substituting it by a parameter $\gamma$ [4] [5]. For instance, we could use the Poveda's approach that considers $\gamma$ as the average value of the operator $\gamma$ in the quantum state $\Psi_{\pm}$ [5]:

$$\hat{\gamma}\to\gamma_{\pm}=\left\langle\Psi_{\pm}\left|\sqrt{1+\frac{\hat{\boldsymbol{p}}^2}{m^2c^2}}\right|\Psi_{\pm}\right\rangle.\tag{12}$$

Or consider, as originally was done by Grave de Peralta, that $\gamma$ is just the relativity Lorentz factor associated to the corresponding classical particle [8]-[15]. Anyway, after substituting the operator $\gamma$ by the parameter $\gamma$, we can use Equation (2) for rewriting Equation (11) as Equation (1). The case $\mu=\mu_+=(1+\gamma)m/2>0$ have been intensively studied before [4] [5] [8] [9] [10] [14] [15]:

$$i\hbar\frac{\partial}{\partial t}\Psi_{+}=-\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_{+}+V\Psi_{+}.\tag{13}$$

Solving Equation (13) gives the energies $E'_+=K_++V$, with $K_+>0$, and the wavefunctions $\Psi_+$ corresponding to a relativistic spin-0 particle of mass $m$, which is moving in the scalar potential $V$ [4] [5] [8] [9] [10] [14] [15]. We will focus our attention now in the case $\mu=\mu_-=-(1+\gamma)m/2<0$:

$$i\hbar\frac{\partial}{\partial t}\Psi_{-}=\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_{-}+V\Psi_{-}.\tag{14}$$

Solving Equation (14) gives the energies $E'_-=K_-+V$, with $K_-<0$, and the

wavefunctions $\Psi_-$ corresponding to the same relativistic spin-0 particle of mass $m$, which is moving in the scalar potential $V$. It should be noted that if γ is chosen as the relativity Lorentz factor associated to the corresponding classical particle, then the wavefunctions $\Psi_{\mp}$ are the two components of the so-called Klein-Gordon equation in the Schrödinger form [3]. When $V = -eU_C$, where $U_C$ is the Coulomb potential $+Ze/(4\pi\varepsilon_o r)$, $r = |\boldsymbol{r}|$, and $\varepsilon_o$ is the absolute dielectric permittivity of the vacuum, Equation (13) can be used for approximately describing (discounting the spin) the quantum states with $K > 0$ of the electron (the particle) in a Hydrogen-like atom with atomic number $Z$ [14] [15]. On the other hand, Equation (14) can be used for approximately describing (discounting the spin) the quantum states with $K < 0$ of the electron (the particle) in a Hydrogen-like atom with atomic number $Z$ [14] [15]. These equations are:

$$i\hbar\frac{\partial}{\partial t}\Psi_p = -\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_p - eU_C\Psi_p,\ U_C = +\frac{Ze}{4\pi\varepsilon_o r},\ \Psi_p = \Psi_+. \tag{15}$$

And:

$$i\hbar\frac{\partial}{\partial t}\Psi_- = \frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_- - eU_C\Psi_-. \tag{16}$$

This pair of equations (Equations (15) and (16)) resembles the hole theory based on the Klein-Gordon and Dirac's equations [3], where there is a one-to-one relationship between $\Psi_-$ and the positron (antiparticle) wavefunction corresponding to a quantum state with $K > 0$ ($\Psi_a$), which satisfies the following equation [3]:

$$i\hbar\frac{\partial}{\partial t}\Psi_a = -\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_a + eU_C\Psi_a. \tag{17}$$

Nevertheless, the pair of uncoupled equations (Equation (11), Equations (13) and (14), or Equations (15) and (16)) is not equivalent to the Klein-Gordon equation because, unlike the Klein-Gordon and the Dirac equations [3] [21], these pairs of equations describe processes where, first, the number of particles is constant (one) [10], and second, the kinetic energy of the particle or is always positive or is always negative. Therefore, their use is particularly useful for studying relativistic processes with these characteristics.

For instance, the following plane waves are solutions of Equations (13) and (14) for a free ($V = 0$) particle:

$$\begin{aligned}&\Psi_p = \Psi_+ = \mathrm{e}^{\frac{i}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}-E'_+t)},\ \frac{E'_+}{mc^2} = (\gamma-1) > 0;\\ &\Psi_- = \mathrm{e}^{\frac{i}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}-E'_-t)},\ \frac{E'_-}{mc^2} = (1-\gamma) = -\frac{E'_+}{mc^2}.\end{aligned} \tag{18}$$

Note that $\Psi_-$ is not a solution of Equation (17) with $V = 0$ because the energy of the antiparticle should be positive and because, if the particle has a linear momentum $\boldsymbol{p}$, then the linear momentum of the antiparticle should be $-\boldsymbol{p}$. Due to the momentum conservation, if $\Psi_p = \Psi_+$ is given by Equation (18), then $\Psi_a$

should be given by:

$$\Psi_a = \mathrm{e}^{-\frac{i}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{r}+E'_+t)},\ \frac{E'_+}{mc^2} = (\gamma - 1) > 0. \tag{19}$$

So that the free particle and antiparticle travel in opposite directions with the same energy $E'_a = E'_+ = K_+ = -E'_- = -K_-$. The wavefunction of the free antiparticle ($\Psi_a$) can then be obtained from $\Psi_-$ by a charge conjugate operation [3], *i.e.*, first finding the complex conjugate of $\Psi_-$ and then formally substituting $E'_-$ in $(\Psi_-)^*$ by $E'_a = E'_+ = -E'_-$. It is easy to show that this is also true for the stationary states of the Equations (16) and (17) [3], *i.e.*, if:

$$\Psi_-(\boldsymbol{r},t) = \Omega_-(\boldsymbol{r})\mathrm{e}^{-\frac{i}{\hbar}E'_-t},\ \Psi_a(\boldsymbol{r},t) = \Omega_a(\boldsymbol{r})\mathrm{e}^{-\frac{i}{\hbar}E'_at}. \tag{20}$$

are the stationary solutions of Equations (16) and (17), or in general of Equations (13) and (14), respectively, then:

$$E'_a = -E'_- = E'_+ \text{ and } \Omega_a(\boldsymbol{r}) = \Omega^*_-(\boldsymbol{r});\ \text{thus, } \langle\Psi_-|\boldsymbol{r}|\Psi_-\rangle = \langle\Psi_a|\boldsymbol{r}|\Psi_a\rangle. \tag{21}$$

Consequently, Equations (14) and (16) can be used for studying the spatial localization of the antiparticle wavefunction. This is because, due to Equation (21), the spatial localization of $\Psi_-$ is a necessary and sufficient condition for the spatial localization of $\Psi_a$.

It is important to note that $\Psi_+$ and $\Psi_-$ are particle (electronic) states with positive and negative kinetic energy values, respectively, while $\Psi_a$ is an antiparticle (positronic) state with positive kinetic energy. Both particle and antiparticle interact with the same external world represented by $U_C$, which is the same in Equations (15), (16), and (17). Therefore, if $U_C$ can spatially confine the electronic states with negative kinetic energy ($\Psi_-$), then the same $U_C$ can confine $\Psi_a$. The opposite is also true, if $U_C$ cannot confine the electronic states with negative kinetic energy ($\Psi_-$), then the same $U_C$ cannot confine $\Psi_a$. For instance, Equation (17) describes a particle (a positron) moving with $K_{pos} > 0$ while repelled by a matter nucleus with $Z$ protons. Unlike the electron wavefunction, $\Psi_p$ in Equation (15), which is localized around the matter nucleus that attracts the electron, the positron wavefunction, $\Psi_a$ in Equation (17), cannot be localized because the positron is repelled by the matter nucleus. Consequently, the wavefunction corresponding to the electron states with $K_- < 0$, $\Psi_-$ in Equation (16), cannot be localized around the nucleus of the Hydrogen-like atom [21] [22].

Formally, Equation (17) can be obtained from Equation (16) by changing the sign of the kinetic energy term in Equation (16), and then changing the sign of the particle's charge without modifying $U_C$ (because $U_C$ does not depend on the particle's charge but on the external charges $Ze$ that produce $U_C$). This is equivalent to the charge conjugation operation [3]; *i.e.*, Equation (17) is obtained by taking the complex conjugate of both sides of Equation (16). Finally, if the matter nucleus is substituted by the corresponding antimatter one, which concludes the charge-inversion of all the particles and antiparticles in consideration, then $U_C$ changes of sign in Equation (17). This transforms Equation (17) in Equation

(15). Consequently, in correspondence with the C-symmetry of quantum electrodynamic [21] [22], an antimatter Hydrogen-like atom should be indistinguishable from the corresponding matter one. Note that quantum electrodynamics is independently C-symmetric and PT-symmetric; therefore, its C-symmetry implies its CPT symmetry.

## 3. Spatial Localization of the Wavefunctions of Particles and Antiparticles in a Coulomb Potential

An instance, illustrating why the C-symmetry in quantum electrodynamics implies that an antimatter world should be undistinguishable from the matter one, is shown in **Figure 1**. In a Hydrogen-like matter atom (**Figure 1(a)**), the electron wavefunction ($\Psi_+ = \Psi_p$) is spatially localized around the matter nucleus due to the inward acceleration of the electron, which is produced by the attractive electrical force between the electron and the matter nucleus. However, the electron ($\Psi_-$) and positron ($\Psi_a$) wavefunctions are not spatially localized around the nucleus of the Hydrogen-like matter atom.

In contrast, in a Hydrogen-like antimatter atom (**Figure 1(b)**), the positron wavefunction ($\Psi_a$), and thus also the electron wavefunction ($\Psi_-$), are both spatially localized around the antimatter nucleus due to the inward acceleration of the positron, which is produced by the attractive electrical force between the positron and the antimatter nucleus. However, the electron wavefunction ($\Psi_+ = \Psi_p$) is not spatially localized around the nucleus of the Hydrogen-like antimatter atom. Let us now describe the matter and antimatter atoms, but primarily referring to the electron (the particle) and its electronic states $\Psi_+$ and $\Psi_-$ in the Hydrogen-like matter and antimatter atoms.

As illustrated in **Figure 1(a)**, the Coulombic attraction between the matter nucleus and the electron spatially localizes the wavefunction of the electron ($\Psi_+ = \Psi_p$) around the nucleus [6] [7]. In agreement with Equation (15), this happens because the negative potential energy associated to $-eU_C\Psi_+$ in Equation (15) balances the positive kinetic energy associated to [5]:

$$\hat{K}_+\Psi_+ = \frac{\hat{p}^2}{2\mu_+}\Psi_+ = -\frac{\hbar^2}{2\mu_+}\nabla^2\Psi_+, \text{ with } \mu_+ = \frac{1+\gamma}{2}m > 0. \tag{22}$$

The force associated to the nucleus-electron interaction points to the nucleus of the Hydrogen-like atom; therefore, due to the positive effective mass of the electron in the state $\Psi_+$ ($\mu_+ > 0$), this force produces an inward acceleration on the electron (in the state $\Psi_+$) that tends to spatially localize $\Psi_+$. In opposition, the kinetic energy of the electron tends to spread $\Psi_+$.

In contrast with what occurs to $\Psi_+$, and in agreement with Equation (16), the same Coulombic interaction between the matter nucleus and the electron, when the electron is in the state $\Psi_-$, cannot localize the electron wavefunction $\Psi_-$, and thus $\Psi_a$, around the nucleus of the Hydrogen-like matter atom. This is because no balance can be reached between the negative potential energy associated to $-eU_C\Psi_-$ in Equation (16), and the negative kinetic energy associated to:

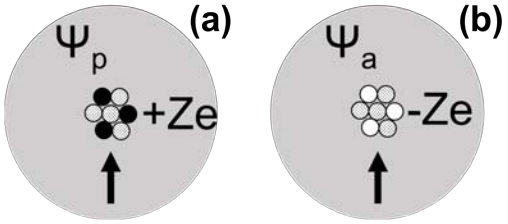



**Figure 1.** Illustration of (a) a matter and (b) antimatter Hydrogen-like atom. The arrows represent the centripetal acceleration of the electron and positron, respectively.

$$\hat{K}_{-}\Psi_{-}=\frac{\hat{p}^{2}}{2\mu_{-}}\Psi_{-}=-\frac{\hbar^{2}}{2\mu_{-}}\nabla^{2}\Psi_{-},\ \text{with}\ \mu_{-}=-\frac{1+\gamma}{2}m<0. \tag{23}$$

As it was stated above, the force associated to the nucleus-electron interaction points to the nucleus of the Hydrogen-like atom; however, when the electron is in the state $\Psi_{-}$, due to the negative effective mass in Equations (16) and (23) ($\mu_{-} < 0$), the same force produces an outward acceleration on the electron that tends to spatially spread $\Psi_{-}$, and thus $\Psi_{a}$. The kinetic energy of the electron in the state $\Psi_{-}$ is negative because $\mu_{-} < 0$, thus large negative values of $K_{-}$ imply large speed values. Consequently, the negative kinetic energy in Equations (16) and (23) still tends to spatially spread $\Psi_{-}$; therefore, it cannot balance the spreading effect of the nucleus-electron interaction on the wavefunction, when the electron is in the state $\Psi_{-}$.

If the matter nucleus of the Hydrogen-like atom were substituted by the nucleus of a Hydrogen-like antimatter atom, then the total charge producing the Coulomb potential would change from $+Ze$ to $-Ze$; therefore, Equations (15) and (16) should be substituted by the following equations:

$$i\hbar\frac{\partial}{\partial t}\Psi_{+}=-\frac{\hbar^{2}}{(\gamma+1)m}\nabla^{2}\Psi_{+}+eU_{C}\Psi_{+}. \tag{24}$$

$$i\hbar\frac{\partial}{\partial t}\Psi_{-}=\frac{\hbar^{2}}{(\gamma+1)m}\nabla^{2}\Psi_{-}+eU_{C}\Psi_{-}. \tag{25}$$

Now, in the antimatter atom, $-U_{C}$ cannot spatially localize the wavefunction of the electron ($\Psi_{p} = \Psi_{+}$) around the antimatter nucleus. In agreement with Equation (24), this would happen because the positive potential energy associated to $+eUC\Psi_{+}$ in Equation (24) cannot balance the positive kinetic energy associated to Equation (22). The force associated to the interaction between the antimatter nucleus and the electron, points in this case away from the antimatter nucleus to the electron; therefore, due to the positive effective mass of the electron in the state $\Psi_{+}$ ($\mu_{+} > 0$), this force produces an outward acceleration on the electron that tends to spatially spread $\Psi_{+}$. Of course, the kinetic energy of the electron also tends to spread $\Psi_{+}$.

As illustrated in **Figure 1(b)**, in contrast with what occurs to the electron wavefunction ($\Psi_{p} = \Psi_{+}$) in a Hydrogen-like antimatter atom, and in agreement

with Equation (25), the same interaction between the electron and the antimatter nucleus could localize the wavefunction $\Psi_-$, and thus, due to Equation (21), the wavefunction of the positron ($\Psi_a$), around the nucleus of a Hydrogen-like antimatter atom. This is because the positive potential energy associated to $+eU_C\Psi_-$ in Equation (25) could balance the negative kinetic energy associated to Equation (23). The force associated to the interaction points away from the antimatter nucleus to the electron; however, when the electron in the antimatter atom is in the state $\Psi-$, due to the negative effective mass in Equation (25) ($\mu_- < 0$), the force produces an inward acceleration on the electron (**Figure 1(b)**) that tends to localize $\Psi_-$, and thus $\Psi_a$. As stated above, the negative kinetic energy in Equations (25) and (23) tends to spatially spread $\Psi_-$; therefore, it can balance the localizing effect of the anti-nucleus-electron interaction, when the electron is in the state $\Psi_-$.

It is worth noting that, in the Hydrogen-like atom, $U_C$ can localize de electronic wavefunction $\Psi_+$ but not $\Psi_-$, while in the Hydrogen-like antiatom $-U_C$ can localize de electronic wavefunction $\Psi_-$ but not $\Psi_+$. Consequently, in agreement with Equation (21), $U_C$ can localize de particle wavefunction ($\Psi_p$) but not the antiparticle one ($\Psi_a$), while $-U_C$ can localize de antiparticle wavefunction ($\Psi_a$) but not the particle one ($\Psi_p$). $U_C$ is charge-inversion-antisymmetric because $U_C$ changes of sign when a C-inversion is applied to the external world surrounding the particle. In addition, the change of sign of $U_C$ is independent of the change of sign of the particle that is moving through $U_C$.

Summarizing the above discussion, we can say the C-symmetry of Equations (13) and (14) indicates that a Hydrogen-like antimatter atom should be as stable as a Hydrogen-like matter atom. If all relevant potentials were of the form $V = \pm eU_C$, with $U_C$ being charge-inversion-antisymmetric, we could confirm and extend the above statement to the whole matter and antimatter. However, as we will discuss in the next Section, we could conceive some relevant potentials that do not have this form.

## 4. Spatial Localization of the Wavefunction of a Single Free Particle or Antiparticle

Gravity potentials are not of the form $V = \pm eU_C$; however, gravity interactions between quantum particles are often weak when compared with electromagnetic interactions. Consequently, we could be tempted to discard the possible influence of gravity potentials on the observed asymmetry between matter and antimatter. However, at least a Newtonian-gravity potential has been previously suggested, as playing an important role in quantum physics [5] [16]. It has been hypothesized that a single free particle could interact gravitationally with itself, due to the spread of its mass density through the extension of its wavefunction [5] [16]. A consequence of this hypothesis, combined with Equation (13), is that no elemental quantum particle with $m > m_P$ could exist (in the state $\Psi_+$) [5]. The gravitational self-interaction potential does not depend on the particle's charge,

and the particle and the antiparticle have the same mass; therefore, one should expect that a similar requirement for existence cannot be obtained using Equation (14) for a particle in the state $\Psi_-$. Indeed, this is the case. From Equation (1), the energy of a particle, which wavefunction, $\Psi_+$ or $\Psi_-$, is localized in a finite space region with radius $r$, could be estimated using the following equation [5]:

$$E'_{\pm}(r) = \frac{\hbar^2}{\mu_{\pm} r^2} - \frac{Gm^2}{r}, \text{ with } \mu_+ = \frac{1+\gamma}{2} m > 0 \text{ and } \mu_- = -\frac{1+\gamma}{2} m < 0. \quad (26)$$

In Equation (26), $G$ is the gravitational constant. The "+" case in Equation (26) was used for obtaining the impossibility of the existence of quantum particles with $m > m_P$ (in the state $\Psi_+$) [5]. This is because $r \to 0$ when $m \to m_P$; *i.e.*, the quantum field ($\Psi_+$) "collapse" when $m = m_P$ [5]. In contrast to $E'_+(r)$, $E'_-(r)$ does not have a local extreme because, first, both the kinetic and potential energies in Equation (26) are negatives when the particle is in the state $\Psi_-$. Second, as illustrated in **Figure 2(b)**, the gravitational self-interaction force points to the "center" of the wavefunction $\Psi_-$; however, due to the negative effective mass in Equations (14) and (23) ($\mu_- < 0$), the gravitational force produces an outward acceleration on the particle in the state $\Psi_-$ that tends to spatially spread $\Psi_-$. For this reason, when the particle is in the state $\Psi_-$, the gravitational force cannot balance the tendency of the kinetic energy to spread $\Psi_-$. This is opposite to what happens for a particle in the state $\Psi_p = \Psi_+$ (Equation (13)). For a particle (in the state $\Psi_+$), as illustrated in **Figure 2(a)**, the same force produces an inward acceleration that tends to localize $\Psi_+$ [5]. Consequently, when the particle is in the state $\Psi_+$, the gravitational force can balance the tendency of the kinetic energy to spatially spread $\Psi_+$.

Elemental quantum particles and antiparticles are created in pairs [3] [17] [21]; thus, an occupied particle state $\Psi_+$ is always created forming a pair with an unoccupied particle state $\Psi_-$ [3] [17] [21]. Consequently, the impossibility of the existence of elemental particles (in the state $\Psi_+$) with $m > m_P$ is a sufficient condition for the impossibility of the existence of elemental particles in the state $\Psi_-$ with $m > m_P$. *I.e.*, neither elemental quantum particles nor antiparticles should exist with $m > m_P$. Indeed, this is the case.

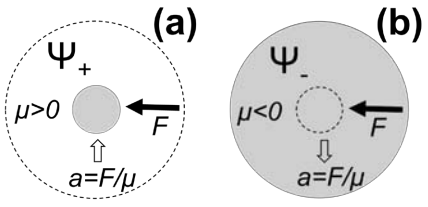



**Figure 2.** Illustration of the spatial (a) localization and (b) spread of the particle wavefunctions $\Psi_+$ and $\Psi_-$, respectively, due to gravitational self-interaction. Solid arrows represent the gravitational force and hollow arrows represent the corresponding acceleration. The discontinuous-line circle indicates the "initial size" of the wavefunctions.

The situation is different for the matter bodies with $m > m_P$ that surround us. These macroscopic objects are formed by numerous quantum particles with mass $m_i < m_P$, $i = 1, 2, \ldots$. It has been suggested that in this case, the collapse of the quantum field ($\Psi_+$), when $m > m_P$, means that these macroscopic objects are really what they look like: classical bodies [5]. Note that macroscopic matter and antimatter objects are not created in pairs; therefore, if macroscopic antimatter objects formed by numerous elemental antiparticles would exist, then from Equation (26) follows they could be massive quantum antimatter objects. This is because no collapse of the corresponding quantum field would occur no matter how massive the antimatter object is.

Self-interaction Newtonian-gravitation and electrostatic potentials are very similar but with opposite signs. The first is attractive, the second is repulsive. This could tempt anyone to explore the consequences of assuming that a single, free, and charged particle could interact electrostatically with itself, due to the spread of its charge density through the extension of its wavefunction. Certainly, this could be considered a controversial hypothesis. However, as it is shown below, the consequences of the electrical self-interaction hypothesis merit the investigation.

As for the gravitational self-interaction, using the electrostatic self-interaction hypothesis in combination with Equations (1), (13), and (14), we obtain that the "size" of a single, free, and charged (with a charge $q$) particle in the states $\Psi_+$ and $\Psi_-$ should be the value of $r$ corresponding to a local extreme of:

$$E'_{\pm}(r) = \frac{\hbar^2}{\mu_{\pm} r^2} + \frac{q^2}{4\pi\varepsilon_0 r}, \text{ with } \mu_+ = \frac{1+\gamma}{2} m > 0 \text{ and } \mu_- = -\frac{1+\gamma}{2} m < 0. \quad (27)$$

The second term in Equation (27) corresponds to the positive potential energy associated to the repulsive electrostatic self-interaction of a particle with itself. Clearly, only $E'_-(r)$ can have a local extreme. This is because the potential and kinetic energies in Equation (27) have different signs if and only if the particle is in the state $\Psi_-$. As illustrated in **Figure 3**, the force associated with the electrostatic self-interaction points away from the "center" of the particle wavefunction. As illustrated in **Figure 3(a)**, this force produces an outward acceleration on the particle that tends to spread $\Psi_+$ (Equations (13) and (27) with $\mu_+ > 0$), but an inward acceleration (**Figure 3(b)**) that tends to localize $\Psi_-$ (Equations (14) and (27) with $\mu_- < 0$). The kinetic energy always tends to spread the wavefunction; consequently, only $\Psi_-$ can be localized by the electrostatic self-interaction. Note that, as this will be discussed in Section 6, the potential responsible of the electrostatic self-interaction is not of the form $V = \pm qU_C$. Equaling to zero the derivative of $E'_-(r)$ respect to $r$, and solving respect to $r$ the resulting equation, we found $E'_-(r)$ has a local maximum when [23]:

$$r = \lambda\!\!\!^{-}_C \xi^{-2} \sqrt{1-\xi^4}, \text{ with } \xi = \frac{q}{q_P}, \; q_P = \sqrt{4\pi\epsilon_0 \hbar c}, \text{ and } \lambda\!\!\!^{-}_C = \frac{\hbar}{mc}. \quad (28)$$

Thus $\lambda\!\!\!^{-}_C \xi^2 \to \lambda\!\!\!^{-}_C$ (the reduced Compton wavelength) when $|q| \to q_P$. Moreover, $r \to 0$, *i.e.*, the quantum field $\Psi_-$ and thus $\Psi_a$ collapse when $|q| \to q_P$. This could

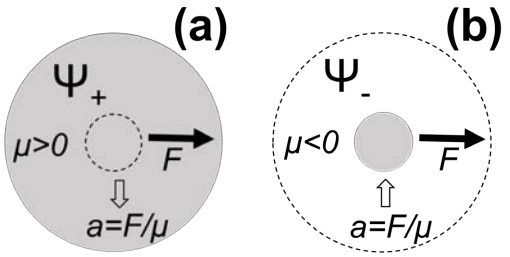



**Figure 3.** Illustration of the spatial (a) spread and (b) localization of the particle wavefunctions $\Psi_+$ and $\Psi_-$, respectively, due to electrostatic self-interaction. Solid arrows represent the electrostatic force and hollow arrows represent the corresponding acceleration. The discontinuous-line circle indicates the "initial size" of the wavefunctions.

be interpreted as the impossibility of the existence of quantum antiparticles with $|q|$ larger than the Plank charge ($q_P \approx 11e$). Indeed, there are no known elemental antiparticles with $|q| > q_P$. This also means that there cannot be elemental particles with $|q| > q_P$ because particles and antiparticles have the same values of $m$ and $|q|$, and they are created simultaneously in pairs.

Following similar arguments than the used for the prediction of the existence of primordial black holes with $m > m_P$ [5] [24], we could now speculate about what could have happened to the elemental free antiparticles with $|q| > q_P$. Or they never existed or, if they existed, they electrically collapsed to a kind of primordial electrical sinks, which are hypothesized for the first time here. If there was a time when no elemental charged antiparticles existed, the primordial charge density fluctuations with $|q| < q_P$ evolved in the elemental charged particles and antiparticles existing today. However, the primordial charge density fluctuations of antimatter with $|q| > q_P$ could have evolved in primordial electrical sinks. Interestingly, if such primordial electrical sinks formed of charged antimatter existed or exist; their existence would imply the existence of an excess of charged matter in the rest of the Universe.

## 5. Absence of Antimatter with a Complicated Structure in the Known Universe

Equations (27) and (28) apply to any quantum object with mass and charge. The quantum field ($\Psi_-$), and thus, due to Equation (21), the antimatter quantum field ($\Psi_a$), both collapse when $|q| > q_P$. This collapse could mean that charged antimatter objects, formed by numerous elemental antiparticles with $|q_i| < q_P$, $i = 1, 2, \ldots$ should be classical antimatter bodies. Note that large matter and antimatter objects are not created in pairs; therefore, this conclusion does not apply to existing charged matter object surrounding us.

Now we can see a possible theoretical explanation for the absence of antimatter with a complicated structure in the known Universe. The existence of molecules,

gases, rocks, planets, oceans, and live in the known Universe requires the existence of atoms. Stable atoms require the interaction of a quantum nucleus with a quantum electron cloud. However, an antimatter nucleus containing more than 11 antiprotons and a cloud containing more than 11 positrons should be classical. Consequently, only very light antimatter atoms should be stable. Indeed, this is both an experimental and an everyday fact [25] [26].

Finally, we should explore the possibility that a charged particle could interact with itself both gravitationally and electrically. This can be accounted for by modifying Equation (27) in the following way:

$$E'_{\pm}(r)=\frac{\hbar^2}{\mu_{\pm}r^2}+\frac{\frac{q^2}{4\pi\epsilon_0}-m^2G}{r},\ \text{with}\ \mu_{+}=\frac{1+\gamma}{2}m>0\ \text{and}\ \mu_{-}=-\frac{1+\gamma}{2}m<0. \quad (29)$$

Therefore, $\Psi_{+}$ can be spatially localized if:

$$\left(\frac{q}{m}\right)^2<4\pi\epsilon_0 G=\left(\frac{q_P}{m_P}\right)^2\approx 7.4\times10^{-21}\frac{\mathrm{C}^2}{\mathrm{kg}^2}. \quad (30)$$

While $\Psi_{-}$ can be spatially localized if:

$$\left(\frac{q}{m}\right)^2>\left(\frac{q_P}{m_P}\right)^2. \quad (31)$$

Note that Equation (29) reduces to Equation (27) when $|q|/m \gg q_P/m_P$. This is the case, for instance, for positrons and antiprotons, that have values of $|q|/m \approx 3.1\times10^{22}$ and $9.2\times10^{15}$ $\mathrm{C}^2/\mathrm{kg}^2$, respectively. For any neutral particle like the neutron, $|q|/m \ll q_P/m_P$, therefore, Equation (29) reduces to Equation (26). From Equation (29) follows that hypothetical original antimatter fluctuations of mass and charge densities such that $|q|/m \gg q_P/m_P$ and $|q|>q_P$, and original matter fluctuations of mass and charge densities with $|q|/m \ll q_P/m_P$ and $m>m_P$ could have collapsed in primordial electrical sinks and gravitational black holes, respectively.

## 6. Discussion

In the previous two Sections was shown that the hypothesis, of a free charged particle interacting electrostatically with itself, seems to imply an asymmetry between matter and antimatter like the one indeed existing in Nature. If a free particle of charge $q$ could interact electrostatically with itself, Equations (13) and (14) should be modified in the following way [16]:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\Psi_p&=-\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_p+V_{si+}\Psi_p,\\ V_{si+}&=\frac{q^2}{4\pi\varepsilon_0}\int\frac{\left|\Psi_{+}(r',t)\right|^2}{|r'-r|}\mathrm{d}V'>0,\ \ \Psi_p=\Psi_{+}. \end{aligned} \quad (32)$$

$$i\hbar\frac{\partial}{\partial t}\Psi_{-}=\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_{-}+V_{si-}\Psi_{-},\ V_{si-}=\frac{q^2}{4\pi\varepsilon_0}\int\frac{\left|\Psi_{-}(r',t)\right|^2}{|r'-r|}\mathrm{d}V'>0. \quad (33)$$

Note that the potential responsible of the electrostatic self-interaction is not of the form $V = \pm qU_C$. This is because $q/(4\pi\varepsilon_o r)$ is not charge-inversion-antisymmetric, *i.e.*, it does not change of sign when a C-inversion is applied to the external world surrounding the particle.. Also, unlike $U_C$, $q/(4\pi\varepsilon_o r)$ does change of sign when the particle changes of sign. As it was discussed in the previous section, $V_{si+}$ in Equation (32) tends to spatially spread $\Psi_p = \Psi_+$. However, the repulsive electrostatic self-interaction tends to spatially shrink $\Psi_-$. Consequently, for the corresponding free antiparticle interacting electrostatically with itself, if Equation (21) should remain valid, Equation (17) should be modified in the following way:

$$i\hbar\frac{\partial}{\partial t}\Psi_a = -\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_a - V_{si-}\Psi_a. \tag{34}$$

So that Equation (34) can be obtained from Equation (33) as Equation (17) can be obtained from Equation (16); *i.e.*, by taking the complex conjugate of both sides of Equation (33). Consequently, the antiparticle should interact with itself in a different way that the charged particle does. Equation (34) differs from Equation (32) in that $-V_{si-}$ does not tend to spatially spread $\Psi_a$ but, like it does to $\Psi_-$, $-V_{si-}$ tends to shrink $\Psi_a$. The discussion, presented in the two previous Sections, is consistent with choosing Equation (34) as the correct equation for the hypothetical self-interacting free antiparticle. This implies that particles interact differently with itself than antiparticles, thus introducing a fundamental asymmetry between elemental charged particles and antiparticles.

From Equation (27) follows that the wavefunction of a free electron or a free proton should be an extended plane wave. However, by evaluating Equation (28) for a positron and an antiproton, we obtain the wavefunction of a free positron should be confined in a spatial region of radius equal to the Bohr radius ($r \approx 0.05$ nm), while the wavefunction of a free antiproton should be confined in a region of radius $r \approx 29$ fm. This may be verifiable predictions of the correctness of Equation (34). Experiments comparing the spectra of matter and antimatter Hydrogen atoms have been reported and continue being conducted [27] [28]. This may provide another possibility for experimentally checking the hypothesis that a single particle could interact with itself, due to the spread of its charge density through the extension of its wavefunction. If the electron could interact electrostatically with itself, Equations (15) and (17) should be modified in the following way:

$$i\hbar\frac{\partial}{\partial t}\Psi_p = -\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_p - \left(eU_C - V_{si+}\right)\Psi_p. \tag{35}$$

$$i\hbar\frac{\partial}{\partial t}\Psi_a = -\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_a - \left(eU_C + V_{si-}\right)\Psi_a. \tag{36}$$

The self-interaction term in Equation (35) is not included in a description of a matter Hydrogen-like atom using the Klein-Gordon or Dirac equations [3] [17].

$V_{si+}$ tends to spatially spread $\Psi_p$, due to the screening of $U_C$ by the electron wavefunction. This is in reminiscence of the Lamb shift, which is explained, in a quantum field theory context, as produced by the screening of $U_C$ by the random fluctuations of the vacuum, thus lasting producing a slight spatially spreading of $\Psi_p$ [29] [30] [31]. In contrast, $-V_{si-}$ tends to spatially shrink $\Psi_a$. Consequently, a small difference between the precise energy level structure of the matter and antimatter Hydrogen atoms should exist. However, if these subtle differences were experimentally discarded, this would not imply that particles could not interact with itself. Such negative experimental result could just mean that Equation (34) should be substituted by:

$$i\hbar\frac{\partial}{\partial t}\Psi_a = -\frac{\hbar^2}{(\gamma+1)m}\nabla^2\Psi_a + V_{si-}\Psi_a. \tag{37}$$

So that Equation (37) can be obtained from Equation (33) by changing the sign of the kinetic energy term in Equation (33) without changing the sign of the potential term. This is not equivalent to taking the complex conjugate of both sides of Equation (33).

However, if the correct equation for a free self-interacting antiparticle were not Equation (34) but Equation (37), then Equation (37) could still be obtained from Equation (33) by changing the sign of the kinetic energy term in Equation (33), and then changing the sign of the particle's charge. This is because in contrast to $-eU_C$ in Equation (16), $V_{si-}$ is proportional to the square of the particle's charge. Also, in contrast to $-eU_C$ in Equation (17), $V_{si-}$ in Equation (37) does not change of sign if the matter nucleus is substituted by the corresponding antimatter one. Consequently, if the correct equation for a free self-interacting antiparticle were not Equation (34) but Equation (37), then particles and antiparticles would interact equally with itself, and matter and antimatter would continue being theoretically symmetric. However, Equation (21) would not be valid for self-interacting potentials. Adopting Equation (37) as the correct antiparticle equation could be justified because $V_{si+}$ and $V_{si-}$ are not, strictly speaking, scalar potentials due to their dependence on $\Psi_+$ and $\Psi_-$. Consequently, Equations (32) and (33) do not strictly follow from Equations (13) and (14) but they really are an *ansatz*.

It should also be noted that the spin of the particle has not been included in the previous discussions. This is because the kinetic energy of a particle, which is in the ground state and spatially confined in a cube of size $r$, does not dependent of the particle's spin and is proportional to $\hbar^2/\mu r^2$ [3] [5] [6] [11], which is the kinetic energy expression used in Equations (26), (27), and (29).

## 7. Conclusions

A pair of decoupled Schrödinger-like, but relativistic quantum mechanics gGP equations were explored. One, with effective mass $\mu_+ > 0$, directly describes a single relativistic quantum particle in a quantum state with $K > 0$. The other, with $\mu_- < 0$, indirectly describes a single relativistic quantum antiparticle in a quantum

state with $K > 0$. The simplicity and the independence of these equations facilitated the study of the spatial localization of the wavefunctions of particles and antiparticles.

It was studied, for the first time, a gGP equation with a negative effective mass value (Equation (1) with $\mu_- < 0$). The solutions of this equation are the wavefunctions $\Psi_-$, which are related to the antiparticle's wavefunctions ($\Psi_a$) through Equations (20) and (21). The energy values corresponding to $\Psi_-$ are $E'_- = -E'_a$, where $E'_a$ is the energy of the associated antiparticle in a quantum state with $K > 0$.

It was found that the antiparticle's wavefunction corresponding to a quantum state with $K > 0$, $\Psi_a$, can be spatially localized by a scalar potential of the form $V = \pm eU_C$ if, and only if, the related particle's wavefunction corresponding to a quantum state with $K < 0$, $\Psi_-$, can be spatially localized (Equation (21)).

As expected, it was found that matter and antimatter are indistinguishable if $U_C$ in Equations (15), (16), and (17) is Coulombic and produced by the external world where the particle or antiparticle exists. When this happens, $U_C$ change of sign under a C-transformation is applied to the external world where the particle exists, but it does not change of sign when the sign of the particle's charge changes. It was shown that in this case, both gGP equations are C-symmetric, and matter and antimatter are undistinguishable.

However, if the potential is not produced by the external world surrounding the particle (or antiparticle) but by the interaction of the particle (or antiparticle) with itself, and Equations (21) and (34) are valid, then matter and antimatter could be distinguished because both gGP equations are not C-symmetric.

Finally, it was shown that if a quantum particle (or antiparticle) could interact with itself, due to the spread of the charge and mass densities through its wavefunction, and Equations (21) and (34) are valid, then there should not be elemental particles and antiparticles with $m > m_P$ or $|q| > q_P$. It was proposed the possible existence of primordial antimatter electrical sinks. These hypothetical antimatter objects could be partially responsible for the observed excess of charged matter in Nature. It was also suggested that all macroscopic matter bodies such that $|q|/m \ll q_P/m_P$ and $m > m_P$, and all macroscopic antimatter bodies such that $|q|/m \gg q_P/m_P$ and $|q| > q_P$, should be classical objects. This could explain the absence of antimatter with a complicated structure in the known Universe.

The author is aware that the results presented in this work could be used in two opposite ways. Those convinced that elemental particles are mathematical points, and convinced of the universality of the CPT symmetry, could argue that this work reinforces their belief. Others, that are perplexed by the fact that in contrast with these beliefs, Nature seems to be mostly made of matter, could argue that this work points to a plausible explanation of why Nature is as it seems to be.

## Conflicts of Interest

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

## References


[1] Canetti, L., Drewes, M. and Shaposhnikov, M. (2012) *New Journal of Physics*, **14**, Article ID: 095012. https://doi.org/10.1088/1367-2630/14/9/095012

[2] Sather, E. (1996) The Mystery of the Matter Asymmetry, Beam Line, Spring/Summer 31, USA.

[3] Greiner, W. (1990) Relativistic Quantum Mechanics: Wave Equations. Spring-Verlag, New York. https://doi.org/10.1007/978-3-662-02634-2

[4] Poveda, L.A., Grave de Peralta, L., Pittman, J. and Poirier, B. (2022) *Foundations of Physics*, **52**, 29. https://doi.org/10.1007/s10701-022-00541-5

[5] Grave de Peralta, L., Poveda, L.A. and Poirier, B. (2021) *European Journal of Physics*, **42**, Article ID: 055404. https://doi.org/10.1088/1361-6404/ac0ecc

[6] Davydov, A.S. (1965) Quantum Mechanics. Pergamon Press, Oxford.

[7] Griffiths, J. (1995) Introduction to Quantum Mechanics. Prentice Hall, Hoboken. https://doi.org/10.1063/1.2808172

[8] Grave de Peralta, L. (2020) *European Journal of Physics*, **41**, Article ID: 065404. https://doi.org/10.1088/1361-6404/aba7dc

[9] Grave de Peralta, L. (2020) *Results in Physics*, **18**, Article ID: 103318. https://doi.org/10.1016/j.rinp.2020.103318

[10] Grave de Peralta, L (2020) *Journal of Modern Physics*, **11**, 196. https://doi.org/10.4236/jmp.2020.112012

[11] López-Boada, R. and Grave de Peralta, L. (2021) *Journal of Modern Physics*, **12**, 1966. https://doi.org/10.4236/jmp.2021.1214113

[12] Grave de Peralta, L. and Farooq, H. (2021) *Journal of Modern Physics*, **12**, 1145-1159. https://doi.org/10.4236/jmp.2021.128068

[13] Ruiz-Columbié, A., Farooq, H. and Grave de Peralta, L. (2021) *Journal of Modern Physics*, **12**, 1358. https://doi.org/10.4236/jmp.2021.1210085

[14] Grave de Peralta, L. (2020) *Scientific Reports*, **10**, Article No. 14925. https://doi.org/10.1038/s41598-020-71505-w

[15] Grave de Peralta, L. (2020) *Journal of Modern Physics*, **11**, 788. https://doi.org/10.4236/jmp.2020.116051

[16] Diósi, L. (1984) *Physics Letters A*, **105**, 199-202. https://doi.org/10.1016/0375-9601(84)90397-9

[17] Strange, P. (1998) Relativistic Quantum Mechanics: With Applications in Condensed Matter and Atomic Physics. Cambridge University Press, New York. https://doi.org/10.1017/CBO9780511622755

[18] Jackson, J.D. (1975) Classical Electrodynamics. 2nd Edition, John Wiley & Sons, New York.

[19] Durand, B. and Durand, L. (1984) *Physical Review D*, **30**, 1904. https://doi.org/10.1103/PhysRevD.30.1904

[20] Jacobs, S., Olsson, M.G. and Suchyta III, C. (1986) *Physical Review D*, **33**, 3338-3348. https://doi.org/10.1103/PhysRevD.33.3338

[21] Dirac, P.A.M. (1930) The Principles of Quantum Mechanics. International Series of Monographs on Physics, Clarendon Press, Oxford.

[22] Salman, M. and Saue, T. (2020) *Symmetry*, **12**, 1121. https://doi.org/10.3390/sym12071121

[23] Poveda, L.A. (2021) Private Communication. Departemento de Fısica, Centro Fed-

eral de Educação Tecnológica de Minas Gerais, Amazonas 5253, 30421-169 Belo Horizonte, MG, Brasil.

[24] Hawking, S. (1971) *Monthly Notices of the Royal Astronomical Society*, **152**, 75-78. https://doi.org/10.1093/mnras/152.1.75

[25] Sellner, S., *et al.* (2017) *New Journal of Physics*, **19**, Article ID: 083023. https://doi.org/10.1088/1367-2630/aa7e73

[26] The STAR Collaboration (2011) *Nature*, **473**, 353-356. https://doi.org/10.1038/nature10079

[27] Andresen, G., Ashkezari, M., Baquero-Ruiz, M., *et al.* (2010) *Nature*, **468**, 673-676. https://doi.org/10.1038/nature09610

[28] The ALPHA Collaboration (2020) *Nature*, **578**, 375-380.

[29] Scully, M.O. and Zubairi, M.S. (1997) Quantum Optics. Cambridge University Press, Cambridge.

[30] Bethe, H.A. (1947) *Physical Review*, **72**, 339-341. https://doi.org/10.1103/PhysRev.72.339

[31] Kroll, N.M. and Lamb, W.E. (1949) *Physical Review*, **75**, 388-398. https://doi.org/10.1103/PhysRev.75.388

Scientific Research Publishing

Journal of Modern Physics, 2022, 13, 1117-1127
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# A Novel Classical Model of the Free Electron

**Arlen Young**

Independent Researcher, Palo Alto, CA, USA
Email: arlen_young@yahoo.com

**How to cite this paper:** Young, A. (2022) A Novel Classical Model of the Free Electron. *Journal of Modern Physics*, **13**, 1117-1127.
https://doi.org/10.4236/jmp.2022.137064

**Received:** June 21, 2022
**Accepted:** July 15, 2022
**Published:** July 18, 2022



## Abstract

Previous models of the free electron using classical physics equations have predicted attributes that are inconsistent with the experimentally observed attributes. For example, the magnetic moment has been calculated for the observed spinning electric charge. For the calculated moment to equal the observed moment, the electron would either have to spin at two hundred times the speed of light or have a charge radius two hundred times greater than the classical radius. A similar inconsistency results when the mass derived from the spin angular momentum is compared with the observed mass. A classical model is herein proposed which eliminates the magnetic moment inconsistency and also predicts the radius of the electron. The novel feature of the model is the replacement of a single charge with two opposite charges, one on the outer surface of the electron and the other at the center.

## Keywords

Classical Electron Model, Free Electron, Electron Structure, Electron Charge, Electron Radius, Electron Spin, Electron Shape, Electron Compressibility

## 1. Introduction

Some attributes of the electron that have been measured are charge, mass, angular momentum, and magnet moment. Angular momentum has been assumed to result from the spinning of the mass. Magnetic moment has been assumed to result from the spinning of the charge. Previous classical models of the electron have attempted to relate these attributes using classical physics equations. The result has been inconsistent on the order of two orders of magnitude. Attempts to resolve the inconsistencies have predicted very large radii or rotation speeds greatly exceeding the speed of light. As a consequence, many have concluded that the classical laws of physics do not apply in the quantum domain of the electron.

The following article proposes a novel model of the free electron using classic-

al physics equations. The model has the following features:

- replaces the single charge in previous models with two opposite charges, one on the outer surface of the core and one at the center;
- eliminates the inconsistency between the observed charge radius and the charge radius deduced from the spin magnetic moment;
- assumes a spin rotation speed that is close to, but does not exceed, the speed of light;
- predicts a radius that is close to the calculated classical and experimentally measured radii;
- suggests that the core shape might be a ring rather than a sphere;
- does not rely on tensile strength of the electron material to hold it together;
- there is no compression or tensile force on the electron core material at the equator;
- predicts a force holding the electron together greater than the nuclear Strong Force.

Except where otherwise noted, all constants, such as those in Table 1, and equations in this article are expressed in cgs units.

## 2. Magnetic Moment and Spinning Charge

### 2.1. Background

Consider a model of the electron wherein the charge $q$ is distributed across the surface of a sphere of radius $R$. Assume the sphere rotates very near the speed of light $c$.

The magnetic dipole moment $M$ of a spinning charged spherical shell is:

$$M = \frac{q}{3}\omega r^2 \quad \text{[MKS] [2]} \quad M = \frac{q}{3c}\omega r^2 \quad \text{[cgs]}$$

where $q$ = uniformly distributed charge;

$r$ = radius;

$\omega = \frac{2\pi}{T}$, where $T$ = period of rotation.

Table 1. Electron constants.

| *constant* | *symbol* | *value* [*cgs*] |
|---|---|---|
| charge | $q$ | $-4.8032 \times 10^{-10}$ |
| mass | $m$ | $9.1094 \times 10^{-28}$ |
| classical radius | $R$ | $2.82 \times 10^{-13}$ |
| spin angular momentum | $S$ | $9.1329 \times 10^{-28}$ |
| magnetic dipole moment | $M$ | $-9.284764 \times 10^{-21}$ |
| Planck's constant | $h$ | $6.6261 \times 10^{-27}$ |
| speed of light | $c$ | $2.99792458 \times 10^{10}$ |

These constants were measured or derived from experimental observations [1]. Spin angular momentum S was derived from the equation $S = \sqrt{3}\frac{h}{4\pi}$.

Assuming the electron model has a spherically charged shell of radius $R$, its spin magnetic moment $M$ can be expressed as:

$$M = \frac{2\pi q}{3Tc} R^2$$

$$\text{rotation speed at the equator} = \frac{2\pi R}{T} = \frac{3Mc}{qR} = 205.6c$$

To generate the observed magnetic moment by spinning the observed charge, the electron equator would have to spin at more than 200 times the speed of light. Since mass cannot spin faster than the speed of light, an alternative explanation for the large observed spin magnet moment might be a radius larger than $R$. Now assume that the rotation speed is less than but very close to the speed of light $c$. The required radius $r$ can be calculated as follows:

$$M = \frac{2\pi q}{3Tc} r^2,$$

where

$$T = \frac{2\pi}{c} r$$

$$r = \frac{3M}{q} = 5.799 \times 10^{-11} = 205.6R$$

The value calculated for $r$ is close to that calculated in [3], which is $3.86 \times 10^{-11}$. The difference could be attributed to [3] assuming the charge is concentrated in a ring, rather than distributed across a sphere.

The classical model of an electron with a spherical charge shell predicts a spin rotation speed of more than 200 times the speed of light. Or, if the spin speed is limited to the speed of light, the radius of the shell would be more than 200 times the classical electron radius. If the charge is assumed to be uniformly distributed throughout the interior of the sphere, the speed or charge shell radius would be even greater. Such large inconsistencies have caused many to believe that classical mechanics and electrodynamics cannot be used to model the electron.

## 2.2. Proposed Charge Model

A model is proposed wherein the electron is comprised of two opposite charges. The spinning outer charge $q^+$ creates the observed magnetic moment. The inner charge $q^-$ located at the center has a very small radius, such that it does not significantly contribute to the magnetic moment. The electric fields from the inner and outer charges combine such that the net electric field of the electron appears to be created by a negative charge $q$ of the observed value.

$$q = q^+ + q^-$$

The cgs unit for charge is $cm^{3/2}\ g^{1/2} \cdot s^{-1}$. Length, mass, and time all change on a speeding platform relative to a stationary platform according to Einstein's Spe-

cial Relativity equations. However, when combined within the unit for charge, the three changes all cancel each other out. Therefore, charge is invariant under speed. A value of charge is the same whether observed on a stationary platform or a platform moving near the speed of light [4].

Magnetic dipole moment $M$ of a spinning charged spherical shell:

$$M = \frac{q^+}{3}\frac{\omega}{c}r^2,$$

where $q^+$ = positive charge;

$r$ = radius;

$\omega = \frac{2\pi}{T}$, where $T$ = period of rotation $= \frac{2\pi}{c}r$.

For sign consistency, the value of $M$ for the electron is considered to be negative, corresponding to a negative charge. Since $M$ is actually being generated in the dual-charge model by a spinning positive charge, instead of a negative charge in the single-charge model, the spin direction in the dual-charge model must be reversed from that in the single-charge model. $M$ after reversing the spin rotation:

$$M = -\frac{q^+}{3}r$$

For $M$ = electron magnetic dipole moment:

$$q^+ = -\frac{3M}{r}$$

$$q^- = q - q^+ = q + \frac{3M}{r}$$

## 3. Mass and Spin Angular Momentum

The angular momentum of a rotating ring is

$$S = mrv,$$

where $m$ = mass of the ring;

$v$ = speed of rotation;

$r$ = radius.

The speed of rotation of the ring is:

$$v = \frac{S}{mr}$$

Let the ring have an angular momentum equal to the spin angular momentum $S$ of the electron, a mass $m$ equal to that of the electron and a radius r equal to that of its classical radius $R$:

$$v = 3.555 \times 10^{12} = 118.6c$$

Special Relativity tells us that the speed of a mass must be less than that of light. Limiting the ring speed $v$ to slightly less than the speed of light $c$, the ring mass $m'$ would have to be slightly greater than:

$$m' = \frac{S}{rc}$$

Consider a non-rotating sphere having a uniformly distributed mass throughout its volume. Slice the sphere into many concentric cylinders, each having a mass $m_n$. Now spin the sphere about the cylinder axes at a speed $v$. The relativistic mass $m_n'$ of each cylinder spinning at a speed $v_n$ is given by the Special Relativity equation:

$$m_n' = \frac{m_n}{\sqrt{1-\left(\frac{v_n}{c}\right)^2}}$$

The total mass $m'$ of the rotating sphere is $m' = \sum_n m_n'$.

An approximation using ten cylinders and a rotation speed of the outer cylinder very near the speed of light shows that almost all of the mass $m'$ is concentrated in the outer cylinder, or ring, to within about 1%. Therefore, for the following calculations, the relativistic mass $m'$ of a rotating sphere will be considered to be uniformly concentrated along a ring of radius $r$ at the equator. For the purposes of modeling the electron, the equation $m' = \frac{S}{rc}$ will be assumed to be a very close approximation for $m'$.

## 4. Radius

An electron can be modeled as originating from a spherical shell of charge $q^+$ having a very large radius. At the center of the shell is a charge $q^-$. The electron radius to be calculated is $R'$. The two charges are:

$$q^+ = -\frac{3M}{R'}$$

$$q^- = q + \frac{3M}{R'}$$

The charge increments on the shell tend to repel each other. Each increment at distance $r$ from the center sees the remaining increments as a point charge at the center.

Coulomb's law for force $f$ between two charges separated by a distance $r$: $f = \frac{q^1 q^2}{r^2}$.

The total repulsive force for all electron charge increments on the spherical shell is:

$$\left(\frac{q^+}{r}\right)^2 = \left(\frac{3M}{R'}\right)^2 \frac{1}{r^2}$$

The attractive force between $q^+$ and $q^-$ is:

$$\frac{q^+ q^-}{r^2} = -\left[\frac{3Mq}{R'} + \left(\frac{3M}{R'}\right)^2\right]\frac{1}{r^2}$$

The sum of the repulsive and attractive forces is a net inward force of:

$$-F = -\frac{3Mq}{R'}\frac{1}{r^2}$$

The outer charge shell will collapse under the inward force $F$. The electrostatic potential energy lost when the shell's radius contracts from infinity to $r$ is:

$$E = -\int_{\infty}^{r}\frac{3Mq}{R'}\frac{1}{x^2}\mathrm{d}x = \frac{3Mq}{R'}\frac{1}{r}$$

The energy $E$ lost is transferred to the spinning electron energy $E'$:

$$E' = m'c^2 = \frac{Sc}{R'}$$

The electron radius $R'$ is the solution $r$ to the equation $E' = E$.

$$R' = \frac{3Mq}{Sc} = 4.8864\times10^{-13} = 1.73R$$

The radius for the proposed dual-charge electron model with a spherical core is 73% greater than the classical radius $R$.

## 5. Internal Forces

In this section, the internal forces of the electron model are calculated as a function of radius $r$, given that charge $q$, magnetic moment $M$, and spin angular momentum $S$ are constants.

The outward forces tending to push the electron apart are:

- centrifugal force on the spin mass ring associated with the spin angular moment $S$;
- mutual repulsion of the charge increments on the outside charge shell;
- outward repulsion of the charge moving through its own magnetic field;
- compression force of the electron core material.

The inward force is the attraction of the spherical charge shell to the opposite charge at the center of the electron.

The model assumes that the spin speed $v$ at the electron equator is slightly less than the speed of light $c$:

$$\text{centrifugal force} = \frac{m'v^2}{r} \cong \frac{m'c^2}{r} = \frac{Sc}{r^2}$$

$$\text{mutual repulsion of the outer charge shell} = \frac{(3M)^2}{r^4}$$

magnetic repulsion of outer charge shell—

A charge increment on the outer shell spins through the magnet field created by all of the other spinning charge increments. To simplify calculations, the spinning sphere was approximated by a spinning ring at the equator of the electron. To calculate the magnetic force on the ring, the ring was split into two rings very close to each other. Each ring had one half the total charge. Each ring spins in the magnetic field of the other. The two rings attract each other. The force on each ring was calculated. The net force on the ring pair was then calculated. By comparison, the magnetic outward force was found to be about ten

thousand times weaker that the electric outward force. It is therefore not significant in the following force calculations.

$$\text{inward force} = \frac{q^+ q^-}{r^2} = -\frac{3Mq}{r^3} - \frac{(3M)^2}{r^4}$$

Note that positive forces are repulsive (outward) and negative forces are attractive (inward).

The total internal force $F$ is the sum of the centrifugal, repulsive, and inward forces:

$$-F = \frac{Sc}{r^2} - \frac{3Mq}{r^3}$$

For $r = R' = \frac{3Mq}{Sc}$, the sum of all the internal forces $F$ is zero.

For $r = R'$, the internal forces are balanced. Unfortunately, the balance is unstable. A small change in radius will cause the force to increase such that it causes a greater change in radius, and so forth. The forces within the electron must always be balanced for it to have stable attributes, such as radius.

A stable internal force balance can be achieved by introducing an incompressible or compressible core. An inward force—$F$ will be counteracted by an outward force $F'$ from the compressed core. The force balance will be stable when

$$\frac{\mathrm{d}F'}{\mathrm{d}r} \geq \frac{\mathrm{d}F}{\mathrm{d}r}.$$

For $\frac{\mathrm{d}F'}{\mathrm{d}r} > \frac{\mathrm{d}F}{\mathrm{d}r}$, a small decrease in radius will cause the outward force $F'$ from the core to increase more than the inward force $F$, resisting the change in radius. A small increase in radius will cause the outward force $F'$ to decrease more than the inward force $F$. The net force change will be inward, resisting the radius change.

For $\frac{\mathrm{d}F'}{\mathrm{d}r} = \frac{\mathrm{d}F}{\mathrm{d}r}$, a small change in radius will cause the outward and inward forces to change by the same amount, so the net change in force will be zero.

For $\frac{\mathrm{d}F'}{\mathrm{d}r} < \frac{\mathrm{d}F}{\mathrm{d}r}$, a small decrease in radius will cause the net inward force $F$ to increase more than the resisting compression force $F'$, resulting in a net force imbalance.

Up to this point, the model of the electron core has been spherical and consisting of an incompressible material. The force components at the surface of the sphere are not uniform in magnitude. The electrical forces are, but the centrifugal force is not. The centrifugal force is greatest at the equator and decreases very rapidly away from the equator. It is zero along the spin axis. The rapid decrease is mainly due to the concentration of mass $m'$ at the equator.

For an incompressible core, it is obvious that $\frac{\mathrm{d}F'}{\mathrm{d}r} > \frac{\mathrm{d}F}{\mathrm{d}r}$. The total force bal-

ance is stable.

The stability of a compressible core is considered in the following:

The compressibility constant $K$ for a sphere is defined by:

$$K = -\frac{1}{V}\frac{\mathrm{d}V}{\mathrm{d}P}, \text{ where } V = \text{volume} = \frac{4\pi}{3}r^3$$

$$P = \text{pressure} = \frac{F'}{A}, \text{ where}$$

$F'$ = total force on the area $A$

$$A = 4\pi r^2$$

$$\mathrm{d}V = \left(4\pi r^2\right)\mathrm{d}r$$

$$\mathrm{d}P = -\frac{1}{K}\frac{1}{V}\mathrm{d}V = \frac{\mathrm{d}F'}{A}$$

$$\frac{\mathrm{d}F'}{\mathrm{d}r} = -\frac{12\pi}{K}r$$

$$\frac{\mathrm{d}F}{\mathrm{d}r} = \frac{2Sc}{r^3} - \frac{9Mq}{r^4} \quad \text{near the equator}$$

$$\frac{\mathrm{d}F}{\mathrm{d}r} = -\frac{9Mq}{r^4} \quad \text{away from the equator}$$

For a stable force balance, $\frac{\mathrm{d}F'}{\mathrm{d}r} \geq \frac{\mathrm{d}F}{\mathrm{d}r}$:

$$\frac{12\pi}{K}R' \geq \frac{9Mq}{R'^4} - \frac{2Sc}{R'^3} \quad \text{near the equator}$$

$$\frac{12\pi}{K}R' \geq \frac{9Mq}{R'^4} \quad \text{away from the equator}$$

## 6. Electron Shape, Size, and Charges

The modeled shape of the electron is a function of the compressibility of the electron material. For incompressible material, $K = 0$ and the shape can be spherical. The upper limits to the values of $K$ for electron radius $R'$ are approximately:

$$K = \frac{12\pi R'^5}{9Mq - 2ScR'} = 7.85\times10^{-32} \quad \text{near the equator}$$

$$K = \frac{4\pi R'^5}{3Mq} = 2.62\times10^{-32} \quad \text{away from the equator}$$

For $K > 0$, the core is compressible and not a perfect sphere. The shape will tend toward that of a ring. It will bulge outward at the equator. For $K > \frac{4\pi R'^5}{3Mq} = 2.62\times10^{-32}$, the core will be collapsed along its spin axis. As $K$ increases further, the core will become a ring spinning around the axis.

All of the above equations containing the constant M assume M was calculated for a sphere.

The magnetic moment for a spinning ring is

$$M = \frac{q}{2}\omega R^2 \quad \text{[MKS] [5]} \quad M = \frac{q}{2c}\omega R^2 \quad \text{[cgs]}$$

Therefore, for a spinning ring, $M$ in the equations assuming a spherical core must be replaced by $\frac{2}{3}M$ and the equations recalculated. The equations for the radii of the spherical core and the ring core are:

$$R'(\text{sphere}) = \frac{3Mq}{Sc} = 4.8864\times10^{-13} = 1.73R$$

$$R'(\text{ring}) = \frac{2Mq}{Sc} = 3.2576\times10^{-13} = 1.16R$$

The radius of the ring core is very close to the calculated value $R$ [6] and the approximate value measured by X-ray diffraction [7]. This correlation provides some evidence that the electron core is better modeled as a ring rather than a sphere, as suggested in [7].

The upper limit of $K$ for a ring core of radius $R'$ **is**:

$$K = \frac{6\pi R'^5}{3Mq - ScR'} = 1.5505\times10^{-32}$$

For $K$ greater than the upper limit, the ring core will collapse.

The positive and negative internal charges are the same for both the spherical and ring core shapes, and are:

$$q^+(\text{sphere}) = -\frac{3M}{R'(\text{sphere})} = q^+(\text{ring}) = -\frac{2M}{R'(\text{ring})}$$
$$= 5.7004\times10^{-8} = 118.7q$$

$$q^-(\text{sphere}) = q + \frac{3M}{R'(\text{sphere})} = q^-(\text{ring}) = q + \frac{2M}{R'(\text{ring})}$$
$$= -5.7484\times10^{-8} = 119.7q$$

For the dual-charge model of the electron, the tensile force on the core material is zero. By comparison, the tensile force on the core for the single-charge model is:

$$\frac{Sc}{R^2} + \left(\frac{q}{R}\right)^2 = 3.5\times10^8$$

The internal binding force for the dual-charge model with a ring core is quite large:

$$\frac{2Mq}{(R')^3} + \frac{(2M)^2}{(R')^4} = 3.1\times10^{10}$$

The nuclear Strong Force, which binds the nucleus of atoms together, is: $2.5 \times 10^9$ [8], so the binding force in the dual-charge electron is about ten times stronger than the Strong Force.

## 7. Summary

A model of the electron has been proposed which has two opposite electrical charges. The positive charge $q^+$ resides on the outer surface of the electron. The negative charge $q^-$ resides at the center of the electron. It has a radius small enough so that the spinning negative charge does not significantly contribute to the net magnetic moment. The shape of the electron can be spherical, a ring, or a shape in between the two. The charges are the same for shapes within this range:

$$q^+ = 5.7004\times10^{-8} = 118.7q\,,\quad q^- = -5.7484\times10^{-8} = 119.7q$$

The radius depends on the shape of the electron core:

$$R'(\text{sphere}) = 4.8864\times10^{-13} = 1.73R\,,\quad R'(\text{ring}) = 3.2576\times10^{-13} = 1.16R$$

The radius of the ring core is very close to the calculated and experimental values, suggesting that the electron is better modeled as a ring rather than a sphere.

The single-charge model has a large inconsistency between the observed spin magnet moment and the calculated moment due to the spinning classical charge. Eliminating the inconsistency would require a radius of more than 200 times the classical radius or a spin rotation speed of more than 200 times the speed of light. The proposed dual-charge model eliminates the inconsistency with a radius close to the classical radius and a spin rotation speed slightly less than the speed of light $c$.

Most of the relativistic spinning mass is concentrated in a ring around the electron equator, even for the spherical core shape.

The classical single-charge electron model implicitly depends on great tensile strength of the electron material to hold the electron together. The dual-charge model does not require any tensile strength. The model depends on an incompressible or compressible core to provide a stable internal force balance. For a spherical core, there will be compressive pressure on the core, except at the equator. For a ring core, there is no compressive force. The maximum compressibility constant for stable internal force balance is $K = 1.5508\times10^{-32}$.

The internal binding force that holds the electron ring core together is one order of magnitude greater than the nuclear Strong Force.

## Acknowledgements

The author thanks Thomas Bohan for reviewing the draft of this article and for his helpful suggestions.

## Conflicts of Interest

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

## References

[1] Fischbeck, H.J. and Fischbeck, K.H. (1982) Formulas, Facts, and Constants: for

Students and Professionals in Engineering, Chemistry and Physics. Springer, Berlin Heidelberg. https://doi.org/10.1007/978-3-642-96681-1

[2] Arora, C.L. (2010) Refresher Course in B.Sc. Physics (Vol. I). S. Chand Publishing, New Delhi, 406.

[3] Saglam, M., Bayram, B., Saglam, Z. and Gur, H. (2020) *Journal of Modern Physics*, **11**, 9-15. https://doi.org/10.4236/jmp.2020.111002

[4] Ivezić, T. (1992) *Physics Letters A*, **162**, 96-102. https://doi.org/10.1016/0375-9601(92)90983-S

[5] Halliday, D., Resnick, R. and Krane, K.S. (2010) Physics, Volume 2. 5th Edition. John Wiley & Sons, New York, 819.

[6] Haken, H., Wolf, H.C. and Brewer, W.D. (2005) The Electron. In: *The Physics of Atoms and Quanta. Advanced Texts in Physics.*, Springer, Berlin, Heidelberg, 69.

[7] Compton, A.H. (1919) *Physical Review*, **14**, 20. https://doi.org/10.1103/PhysRev.14.20

[8] Styer, D.F. (2022) Notes on Nuclear and Elementary Particle Physics. https://www2.oberlin.edu/physics/dstyer/Modern/NuclearPhysics.pdf

Journal of Modern Physics, 2022, 13, 1128-1145
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Particle Creation from Yang-Mills Gravity

**Edwin Eugene Klingman**

Cybernetic Micro Systems, Inc., San Gregorio, CA, USA
Email: klingman@geneman.com

**How to cite this paper:** Klingman, E.E. (2022) Particle Creation from Yang-Mills Gravity. *Journal of Modern Physics*, **13**, 1128-1145.
https://doi.org/10.4236/jmp.2022.137065

**Received:** June 28, 2022
**Accepted:** July 24, 2022
**Published:** July 27, 2022



## Abstract

The genesis of physical particles, a foundational aspect of physics, is still a mystery. Quantum field theory creation operators provide an abstract mechanism to bring particles into existence. The assumption of a primordial field underlies the Standard Model (SM), yet the forces have failed to converge to such a field. Current treatments of a superfluid-based universe [Huang, Volovik, and Svistunov, Babaev, Prokof'ev] focus heavily on vortices and Yang-Mills theory, so we analyze self-interaction of the primordial field in the context of Yang-Mills. We show that a self-stabilizing higher-order self-interaction interpretation of the Yang-Mills non-Abelian term yields a stable quantum gravity explanation of the mass-gap. In future we will address the spin-$\frac{1}{2}$ and conserved charge aspects in terms of this fundamental theory of particle creation.

## Keywords

Self-Stabilized Field, Self-Organizing Structure, N$^{th}$-Order Dynamics, Heaviside Equations, Solitons

"*The invention of Yang and Mills was not the first non-Abelian gauge field known to physicists; the gravitational field has that honor.*" Bryce DeWitt.

## 1. Introduction

Particle creation is still a mystery. Since 1954 the Yang-Mills gauge field theory of self-interaction has been believed to be the appropriate framework in which to formulate the problem, but it has so far been impossible to explain the "mass-gap" issue. The *mass-gap* is the finite value of the lowest particle mass above the vacuum energy state. The insurmountability of this problem has inspired a million-dollar *Millennium Prize*, but the prize has been unclaimed for two decades. In this paper I analyze the Yang-Mills formalism and propose a reinterpretation of the non-Abelian self-interaction term that is dynamic in nature. I show that

this interpretation leads to a stable state for particle mass with finite energy above the vacuum state. In the next section I summarize the Quantum Field Theory approach and contrast this with the approach taken herein.

## 2. Quantum Field Theory Approach to Particle Creation

Quantum field theory (QFT) provides a bookkeeping system with symbolic *creation operators* bringing particles into existence while *annihilation operators* subtract particles from the ledger. The operators operate on particle-specific quantum fields. In QFT the quantum fields are more fundamental than the particles, which are viewed as excited states of the fields. For quanta such as photons, the oscillations in the field were viewed as arising from oscillators, which exist at every potential minimum. Zee [1] presents QFT as a mattress, idealized as a 2D lattice of mass points connected to each other by springs, a series of harmonic oscillators. He remarks that, even after a century has passed, the whole subject of QFT remains rooted in this harmonic paradigm; unable to break from the basic notions of oscillations and wave packets. He hopes to get beyond this conception, yet the math formalism fit the oscillator ladder so beautifully with "raising" and "lowering" operators promoted to "creation" and "annihilation" operators. The idea then extended to particles as excited states of quantum fields, with each particle arising from a specific field—the *electron field*, the *muon field*, etc. such that, when Feynman [2] developed a quantum field theory of gravity he treated gravity as the "31st field". Instead of a "*field per particle*"; we assume a "*field per universe*", a primordial field existing at the Big Bang, and ask how the field produces a known particle spectrum based on a "mass gap", or finite energy above the vacuum state. The *Standard Model* assumes all forces converge to such a primordial field, but such has not yet been shown.

Physics is largely based on formulating interactions as changes induced by sources, represented as $\nabla\boldsymbol{\psi} = \boldsymbol{j}$, where $\nabla$ is a change operator that generates changes in the field $\boldsymbol{\psi}$ induced by source $\boldsymbol{j}$, separate from field $\boldsymbol{\psi}$. For primordial field $\boldsymbol{\psi}$ nothing is separate from $\boldsymbol{\psi}$; only field $\boldsymbol{\psi}$ exists. Thus, any change operator operating on field $\boldsymbol{\psi}$ must be equivalent to $\boldsymbol{\psi}$ interacting with itself. This *Self-Interaction Principle* [3] is represented by self-interaction eqn:

$$\nabla\psi = \psi\psi \tag{1}$$

To be meaningful, field $\psi$ and operator $\nabla$ must depend on some variable parameter $\xi$, so we extend our formalism via $\psi \to \psi(\xi)$ and $\nabla \to \partial_\xi$ with two formal solutions—for scalar $\xi$ and for vector $\boldsymbol{\xi}$.

$$\psi(\xi) = -\xi^{-1}, \quad \psi(\boldsymbol{\xi}) = \boldsymbol{\xi}^{-1} \tag{2}$$

We assign physical meaning to these terms; if scalar $\xi$ = time, then $\xi^{-1}$ is frequency; if vector $\boldsymbol{\xi}$ = location in space, then $\boldsymbol{\xi}^{-1}$ is inverse distance. Corresponding operators are $\nabla_t = \partial/\partial t$ and $\nabla_{\boldsymbol{r}} = \partial/\partial\boldsymbol{r}$ so we attempt to solve self-interaction Equation (1). Almeida [4] noted: "*choice of a particular algebra is irrelevant from the point of view of the mathematical validity of the equation*,

*but it may make a significant difference to the perception and comprehension of the physics behind the equation*." If so, the question arises as to the optimal algebra for solution of the self-interaction equation. Einstein and Wheeler viewed physics as geometry, with differential geometry the optimal algebra. Quantum physicists evolved Hilbert-space algebra and group theory symmetry represented by matrix algebra. In 1965 Hestenes evolved Clifford algebra to Geometric Algebra; the only mathematical framework in which every term has both an algebraic and a geometric interpretation [5]. For 3-spatial-dimensions-plus-time the terms include scalars, vectors, bivectors, trivectors, and pseudoscalars, interpreted as duality operators represented by $i$, that transform an entity into its dual. The new relation is geometric product $\boldsymbol{uv}=\boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{u}\wedge\boldsymbol{v}$. Bivector $\boldsymbol{u}\wedge\boldsymbol{v}$ is a directed area representing rotation of $\boldsymbol{u}$ into $\boldsymbol{v}$. Duality operator $i$ transforms this bivector into an axial vector: $\boldsymbol{u}\wedge\boldsymbol{v}=i\boldsymbol{u}\times\boldsymbol{v}$. Substituting the vector derivative for $\boldsymbol{u}$ the geometric product is:

$$\begin{array}{c}\nabla\boldsymbol{v}=\nabla\cdot\boldsymbol{v}+\nabla\wedge\boldsymbol{v}\\ /\quad\quad|\quad\quad\backslash\\ \text{gradient}=\text{div}\;+\;\text{curl}\end{array}\tag{3}$$

No other math formalism has this relation. When $\psi=\boldsymbol{G}(\boldsymbol{r},t)+i\boldsymbol{C}(\boldsymbol{r},t)$ and $\nabla=\boldsymbol{\nabla}+\partial_t$, then Equation (1) takes the form

$$(\boldsymbol{\nabla}+\partial_t)(\boldsymbol{G}+i\boldsymbol{C})=(\boldsymbol{G}+i\boldsymbol{C})(\boldsymbol{G}+i\boldsymbol{C})\tag{4}$$

Expansion of (4) in terms of geometric products and grouping of like terms yields:

| **Self-Interaction equations** | **Heaviside equations** | |
|---|---|---|
| $\boldsymbol{\nabla}\cdot\boldsymbol{G}=\boldsymbol{G}\cdot\boldsymbol{G}-\boldsymbol{C}\cdot\boldsymbol{C}$ | $\boldsymbol{\nabla}\cdot\boldsymbol{G}=-\rho$ | (5a) |
| $i\boldsymbol{\nabla}\cdot\boldsymbol{C}=i2\boldsymbol{G}\cdot\boldsymbol{C}$ | $\boldsymbol{\nabla}\cdot\boldsymbol{C}=0$ | (5b) |
| $\partial_t\boldsymbol{G}-\boldsymbol{\nabla}\times\boldsymbol{C}=\boldsymbol{G}\times\boldsymbol{C}\pm\boldsymbol{C}\times\boldsymbol{G}$ | $\boldsymbol{\nabla}\times\boldsymbol{C}=-\rho\boldsymbol{v}+\partial_t\boldsymbol{G}$ | (5c) |
| $i\boldsymbol{\nabla}\times\boldsymbol{G}+i\partial_t\boldsymbol{C}=0$ | $\boldsymbol{\nabla}\times\boldsymbol{G}=-\partial_t\boldsymbol{C}$ | (5d) |

The equations on the left-hand side of (5) derive from (4) in straightforward fashion. With physical meaning assigned to field $\psi$, one obtains the equations on the right side, derived in 1893 by Heaviside [6], wherein $\boldsymbol{G}$ is gravity and $\boldsymbol{C}$ is the gravitomagnetic field. Decades later the eqns were erroneously labeled the *weak field approximation* to Einstein's non-linear field equations.

Self-interaction Equation (6a) yields the Heaviside-Newton equation. The Poynting-like $\boldsymbol{G}\times\boldsymbol{C}$ terms are momentum density and can be transported in opposite directions, based on initial and boundary conditions imposed locally; hence the ± in (5c); they are represented as $\rho\boldsymbol{v}$ in Heaviside (5c), while field energy density terms, $\boldsymbol{C}\cdot\boldsymbol{C}$ and $\boldsymbol{G}\cdot\boldsymbol{G}$, are represented by $\rho$ in (5a). The time independent gravitational field in (5d) is irrotational, shown by Michaelson-Gale in 1925.

If local field density accelerates, then local gravitomagnetic circulation alters appropriately; the moving density drives the local field. If local density decelerates, change in circulation induces a *gmf*, a *gravito-motive force* $\boldsymbol{F}=-\mathrm{d}\boldsymbol{p}/\mathrm{d}t$

to drive the particle forward. In the vacuum state (the local ether) this Lenz-law-like behavior explains conservation of momentum, which Feynman claimed was inexplicable. In (5b), $\nabla\cdot\boldsymbol{C}=0$, we use of vector identity $\nabla\cdot\nabla\times\boldsymbol{A}=0$ to replace $\boldsymbol{C}$ with a potential vector $\nabla\times\boldsymbol{A}$. Compatible with Equation (5) are the gauge field equations:

$$\boldsymbol{C}=\nabla\times\boldsymbol{A},\ \ \boldsymbol{G}=-\nabla\phi-\partial_t\boldsymbol{A},\ \ \partial_t\phi+\nabla\cdot\boldsymbol{A}=0 \tag{6}$$

The first two Equations in (6) define the fields in terms of the four-potential $A$, while the last equation specifies the Lorenz gauge condition, $\partial_\mu A^\mu=0$. The scalar potential $\phi=-m/r$, and vector potential $\boldsymbol{A}=\boldsymbol{v}$. In analogy with Maxwell's equations, we formulate gauge field four-potential $A=\{\phi,\boldsymbol{A}\}$. Since $\boldsymbol{G}=-\nabla\phi+\partial_t\boldsymbol{A}$ if $\phi$ is constant then $\boldsymbol{G}=\partial_t\boldsymbol{A}$, but since $\boldsymbol{G}$ is the acceleration of gravity, then $\boldsymbol{G}=\mathrm{d}\boldsymbol{v}/\mathrm{d}t\Rightarrow\boldsymbol{A}=\boldsymbol{v}$. Since $\boldsymbol{C}=\nabla\times\boldsymbol{A}$ then $\boldsymbol{C}=\nabla\times\boldsymbol{v}$ is dimensionally correct; $|\boldsymbol{C}|\sim t^{-1}$. With gravitational potential $\phi=-M/r$ the $\boldsymbol{G}$-field has spatial dependence $|\boldsymbol{G}|\sim r^{-2}$; correct for Newtonian mass. For the primordial field, as shown in several of the references, $|\boldsymbol{G}|\sim r^{-1}$. Physically, all Newtonian mass is treated as entirely within the sphere of radius $r$, whereas the mass of the primordial gravitational field is based only on the portion of the field within the sphere. In all cases, with local mass density $\rho$ the interaction energy density of the field is $\boldsymbol{j}\cdot\boldsymbol{A}$ where $\boldsymbol{j}=\rho\boldsymbol{v}$. Heaviside current density $\boldsymbol{j}$ is momentum density $\boldsymbol{p}=\rho\boldsymbol{v}$; the interaction density of the field is $\boldsymbol{p}\cdot\boldsymbol{A}=\boldsymbol{p}\cdot\boldsymbol{v}=\rho v^2$. The field strength matrix constructed from the above [7] is shown:

$$F_{\mu\nu}=\begin{bmatrix}0 & G_x & G_y & G_z\\ G_x & 0 & -C_z & C_y\\ G_y & C_z & 0 & -C_x\\ G_z & -C_y & C_x & 0\end{bmatrix} \tag{7}$$

A full unification of gravitation, electromagnetism, the strong and weak nuclear forces, has not yet been derived. Nevertheless, the four fundamental interactions are generated by a single principle, the gauge principle [8]. Weyl, in 1929, derived the conservation laws and expressed the Riemann tensor in the tetrad form: $R^a_{\mu\nu b}=\left[D_\mu,D_\nu\right]^a_b=\partial_\mu A^a_{\nu b}-\partial_\nu A^a_{\mu b}+A^a_{\mu c}A^c_{\nu b}-A^a_{\nu c}A^c_{\mu b}$. For Yang-Mills, expression of field strength $F_{\mu\nu}=\left[D_\mu,D_\nu\right]$ as commutation was not common at the time; direct expression as a curl was so simple: Weyl's equation is expressed $R=\partial\wedge A+\left[A,A\right]$. Yang stated that, when they presented their theory, they had no idea it might be related to gravitation:

"…*when Mills and I worked on non-Abelian gauge fields, our motivation was completely divorced from general relativity, and we did not appreciate that gauge fields and general relativity are somehow related.*"

Little surprise that, in search of a generalization of isotopic spin for application to the nuclear physics of the "50's, Yang and Mills, as particle physicists, did not have tetradic formulations of general relativity in mind, nor the fiber bundle approach developed through differential forms. Today our preferred framework is Hestenes" Geometric Calculus.

## 3. Aspects of Isospin

Initially Pauli added spin to the Hamiltonian based on energy $\boldsymbol{\mu}\cdot\boldsymbol{B}$ in magnetic field $\boldsymbol{B}$ where magnetic moment $\boldsymbol{\mu}$ is proportional to spin $\boldsymbol{s}$ of the charge, conceived classically. The equation of motion $\dot{\boldsymbol{s}}=\boldsymbol{B}\times\boldsymbol{s}$ results [9] in spin precessing about the $\boldsymbol{B}$-field lines of force in two stable configurations, $\pm\boldsymbol{\mu}\cdot\boldsymbol{B}$. Pauli invented 2 × 2 matrix operator $\hat{\sigma}$ to satisfy $\hat{\sigma}|s\rangle=\pm|s\rangle$ for state $|\boldsymbol{s}\rangle=\begin{pmatrix}up\\dn\end{pmatrix}$, with $\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix}$, $\sigma_y=\begin{pmatrix}0&-i\\i&0\end{pmatrix}$, $\sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$.

Heisenberg conceived of the known nucleons, proton and neutron, as a single particle with two states, ignoring electric charge. Instead of up or down *spin* state he formulated the nucleon state with internal "*isospin*" symmetry to allow Pauli's $\boldsymbol{\sigma}$ matrix to switch between internal symmetry states, $\psi=\begin{pmatrix}\text{proton}\\\text{neutron}\end{pmatrix}$.

Matrices are representations of group symmetry, yet isospin is *not* an exact symmetry; it is only *approximate* since the masses of the proton and neutron are not equal. The matrices $\{\sigma_x,\sigma_y,\sigma_z\}$ represent the 2 × 2 Pauli spin matrices of quantum mechanics. Hestenes constructs an equivalent orthonormal basis of three bivectors $\{\beta_x,\beta_y,\beta_z\}$ satisfying $\beta_x\beta_y=-i\beta_z$. The algebras (with Kronecker delta $\delta_{jk}$ and Levi-Civita alternating symbol $\epsilon_{jkl}$) are written:

| Pauli matrix algebra | Hestenes bivector algebra |
|---|---|
| $\sigma_j\sigma_k=-\delta_{jk}-i\varepsilon_{jkl}\sigma_l$, | $\beta_j\beta_k=-\delta_{jk}-i\varepsilon_{jkl}\beta_l$ (8) |

Bivector algebra is identical to spin matrix algebra, by inspection. Since the algebras are identical, their physical implications should be the same; our expressed preference is for the geometric algebra formulation with geometric elements providing visible structures. Attempts to make gauge fields visible in differential geometry center around fiber bundles, with cartoon-like representations of the type shown in Huang's *Fundamental Forces of Nature* [10]. As Penrose has remarked [11] Yang-Mills isospin fields don't exist in the physical world as far as we know. They are *non-physical* abstractions.

Systems coupled to the electromagnetic field possess global gauge invariance *before the coupling is turned on*, so Schrödinger's equation is invariant under a constant phase change $\psi\to e^{i\alpha}\psi$ where $\alpha$ is constant, since $\partial\left(e^{i\alpha}\psi\right)=e^{i\alpha}\partial(\psi)$. Global phase has no physical consequence. Based on Noether, global gauge invariance guarantees existence of a conserved current that yields charge conservation, and Yang and Mills hoped to find such gauge conservation principles in their treatment of isospin. But the system is <u>not</u> invariant under local transformation $\psi\to e^{i\beta(x)}\psi$ since Schrödinger's equation is not invariant: $\partial\left(e^{i\beta(x)}\psi\right)\neq e^{i\beta(x)}\partial(\psi)$. Global gauge invariance is extended to local gauge invariance by replacing derivative $\partial$ with covariant derivative $D$:

$$\partial\to D,\quad D\to\partial+\frac{iqA}{\hbar}. \tag{9}$$

In quantum mechanics $qA$ is combined with momentum $p$ therefore $qA/\hbar$ has dimension 1/*length*, appropriate to the derivative term. Such derivatives in

physics typically represent translations or rotations in local space, parallel transport along a path. A $U(1)$ rotation through angle $\theta$ can be represented by $e^{i\theta}$ phase factor, which, for infinitesimally small angles, reduces to $1+i\theta$. An arbitrary rotation about a fixed axis can be constructed from successive infinitesimal rotations about that axis. For three axes there are three possible infinitesimal rotations: $1+i\theta_1 L_1$, $1+i\theta_2 L_2$, $1+i\theta_3 L_3$. While the $U(1)$ group of transformations about one axis is Abelian (commuting), continuous transformations about 3 axes form a non-Abelian Lie group satisfying $\left[L_a, L_b\right] = i\varepsilon_{abc} L_c$, (see bivector algebra of Equation (8)). The $L_i$ cannot be numbers since they do not commute. Since any 2 × 2 matrix is a linear combination of the Pauli spin matrices; a generator of rotations about 3 axes can thus be represented $L_a = \sigma_a/2$, with general transformation $U = \exp\left(\frac{i}{2}\omega_a\sigma_a\right)$ where $\omega_a$ are real numbers. The 2 × 2 unitary matrix $U$ has symmetry group $SU(2)$ so isospin is an "internal" symmetry with $SU(2)$ symmetry by construction; operation on any two-component wave function $\psi = \left[\psi_1, \psi_2\right]^{\mathrm{T}}$ with rotation $U$ satisfies $\psi \to U\psi$. In this way the geometry of classical physics is applied to abstract internal symmetry such as isospin.

Yang and Mills, in terms of the infinitesimal charge generator $L_a$ of $SU(2)$, replaced derivative $\partial$ by covariant derivative $D = \partial + \frac{ig}{\hbar} L_a A_a$ in equation of motion $\left(\partial - igA\right)\psi = 0$ where $A_a$ is a 4-vector gauge field with three internal components corresponding to the generation of the gauge group of isospin rotations. $D$ generates a coupling between the particle and the gauge field with interaction energy density $j_a A_a$ where $j_a$ is conserved isotopic spin current density. "*But in the real world, isotopic spin is not conserved; the gauge symmetry is not exact.*" Yang and Mills next guessed that adding quadratic terms to the field strength would represent self-interaction of the gauge field:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu + ig\left[A^\mu, A^\nu\right] \tag{10}$$

Yang-Mills gauge theory is based on an abstract, non-physical, idea of *approximate* isospin symmetry. Yang-Mills theory does <u>not</u> explain the mass gap that is the key to particle physics, so we switch to the *exact* symmetry derived from the fundamental principle of self-interaction:

$\nabla\psi = \psi\psi$ $\Rightarrow$ Heaviside equations $\Rightarrow$ Einstein field equations.

Whereas general relativity is derived from an approximate principle, the *Equivalence Principle*, the Heaviside equations are derived from an exact principle, the *Self-Interaction Principle*. There are several consequences of these facts, treated in [12] [13] [14] [15]. Two key facts: 1) Heaviside theory is equivalent to curved space theory, and 2) Heaviside's equations hold at all scales, from Planck scale to Cosmic Microwave Background.

## 4. Details of Yang-Mills Theory

Yang & Mills [16] formulate $B_\mu$ with 12 independent components: 4 × 4 less

diagonal elements. For a two-component wave function, $\psi$, describing a field with isospin $\frac{1}{2}$, the isotopic gauge transformation $\psi = S\psi'$ where $S$ is a $2 \times 2$ matrix with determinant unity, and all drivatives of $\psi$ appear in combination $\left(\partial_\mu - i\epsilon B_\mu\right)\psi$ where $B_\mu$ are $2 \times 2$ matrices for $\mu = 1,2,3$. Invariance requires

$$S\left(\partial_\mu - i\epsilon B'_\mu\right)\psi' = \left(\partial_\mu - i\epsilon B_\mu\right)\psi .$$

The Yang-Mills isotopic gauge transformation on $B_\mu$, corresponding to $A'_\mu = A_\mu + \frac{1}{e}\frac{\partial \alpha}{\partial x_\mu}$ is

$$B'_\mu = S^{-1}B_\mu S + \frac{i}{\epsilon}S^{-1}\frac{\partial S}{\partial x_\mu} \tag{11}$$

with the last term like the gradient term in the gauge transformation of electromagnetic potentials. To obtain gauge invariant field strengths they define the analog of the electromagnetic case

$$F_{\mu\nu} = \frac{\partial B_\mu}{\partial x_\nu} - \frac{\partial B_\nu}{\partial x_\mu} + i\epsilon\left(B_\mu B_\nu - B_\nu B_\mu\right) \text{ with } F'_{\mu\nu} = S^{-1}F_{\mu\nu}S \tag{12}$$

Yang and Mills next introduce isotopic spin "angular momentum" matrices $\tau^i\left(i = 1,2,3\right)$ which correspond to the isotopic spin of the field $\psi$ under consideration. The $B$ field is then defined as $B_\mu = 2b_\mu \cdot \tau$ where both $b_\mu$ and $\tau$ are 3-component vectors in isotopic space. Interaction with any field $\psi$ of arbitrary isospin requires replacing ordinary derivative of $\psi$ by $\left(\partial_\mu - i\epsilon b_\mu \cdot \tau\right)\psi$ with $\tau$ representing isotopic spin "angular momentum" as above. The isotopic-gauge covariant field strengths $F_{\mu\nu}$ are expressible $F_{\mu\nu} = f_{\mu\nu} \cdot \tau$ where

$$f_{\mu\nu} = \frac{\partial b_\mu}{\partial x_\nu} - \frac{\partial b_\nu}{\partial x_\mu} - 2\epsilon b_\mu \times b_\nu \tag{13}$$

and $f_{\mu\nu}$ transforms like a vector under an isotopic gauge transformation. The field equations derive from the total Lagrangian density

$$\mathcal{L} = \frac{1}{4}f_{\mu\nu} \cdot f_{\mu\nu} . \tag{14}$$

Finally, they define

$$\Im_\mu = J_\mu + 2\epsilon b_\nu \times f_{\mu\nu} \tag{15}$$

with equation of continuity $\partial\Im_\mu / \partial x_\mu = 0$ and the supplementary condition (corresponding to the Lorenz gauge) $\partial b_\mu / \partial x_\mu = 0$ which eliminates the scalar part of the field in $b_\mu$. Equation (15) shows that isotopic spin arises from both spin $\frac{1}{2}$ field ($J_\mu$) and from the $b_\mu$ field itself, thus making the field equations for the $b_\mu$ field nonlinear. This is as far as we will carry Yang and Mills theory in its original form. Writing for the Clay Mathematics Institute, Jaffe and Witten:

"*There is no known way of deriving the mass gap from the original theory.*"

## 5. Angular Momentum and Yang-Mills

Linking to our primordial field $\psi = \boldsymbol{G} + i\boldsymbol{C}$ we identify gravitomagnetic gauge field $\boldsymbol{v}$ with the Yang-Mills $b_\mu$ field. The problem is to create a mass-gap that has evaded physicists since the introduction of the theory. The gravitomagnetic field has also evaded physicist's standard model of particle physics, suggesting a need to reinterpret non-linear fields.

Yang and Mills introduce and discuss isotopic spin "angular momentum" in quotes and are unsure what it means physically. They adapt Pauli's *SU*(2) spin matrices to Heisenberg's *isospin*; a mathematical formalism applied to an abstract *internal* symmetry. The nature of spin, at least classically, is rotation, and rotation in 3D space entails angular momentum. Exactly what is entailed in the space of internal symmetry, represented by gauge field $b_\mu$, is unknown. However, the nature of this gauge field is captured by the curl operation, so it must somehow entail an analog of angular momentum, as Einstein and deHaas [17] showed to be possessed by the magnetic field. Yang and Mills "*define isotopic gauge as an arbitrary way of choosing the orientation of the isotopic spin axis at all space-time points*."

The matrix $S^{-1}\dfrac{\partial S}{\partial x_\mu}$ appearing in Equation (11) is a linear combination of isotopic spin "angular momentum" matrices $\tau^i$ $(i = 1,2,3)$ corresponding to isotopic spin of the field we are considering. The $B_\mu$ matrices contain a linear combination of matrices $B_\mu(x) = \sum_{a=1}^{n} b_\mu^a(x)\tau_a$ or $B_\mu = 2b_\mu \cdot \tau$ where $b_\mu$ and $\tau$ are 3-component vectors in isotopic space. In Heaviside isotopic space, the $b_\mu$ vector is the $v_\mu$ velocity vector determining the linear combination of the bivector angular momenta.

Although there is no well-defined idea of isotopic spin "angular momentum", gravitomagnetic C-field possesses angular momentum; and is proportional to angular momentum: $\boldsymbol{C} = \left(g/c^2\right)\boldsymbol{r} \times \boldsymbol{p}$ with dimension $t^{-1}/l^3$. For $F_{\mu\nu}$ depicted in **Figure 1** we pair $C_y$ with $-C_y$, and cyclical iterations, where the index represents the axis about which these components of the field rotate. In other words, the formalism contains the angular momentum aspect of the components. The C-field components are compatible with the three bivectors shown in the 3-space representation at the right, defined by the *x*, *y*, and *z* axes. *The nature of C-field circulation, from every perspective, is angular momentum.*

Consider Yang-Mills term $\epsilon\left[A_\mu, A_\nu\right]$. The $\epsilon$ corresponds to the *isospin charge* analogous to electric charge *q* that interacts with electromagnetic gauge $A_\mu$ in the Hamiltonian, appearing as $qA_\mu$, the momentum term. For the C-field, $\epsilon$ corresponds to mass, hence $\epsilon\boldsymbol{A} \to m\boldsymbol{v}$, the field momentum (actually $\rho\boldsymbol{v}$ the momentum density). In the original Yang-Mills the $\mu = 1$ term interaction with the $\nu = 2$ term concerns the $mA_xA_y$ term. The geometric algebra product $A_xA_y = A_x \cdot A_y + i\left(A_x \times A_y\right)$. The scalar product vanishes while the curl is proportional to $A_z$. The curl is antisymmetric, so we have

$$A_xA_y - A_yA_x = 2A_z. \tag{16}$$

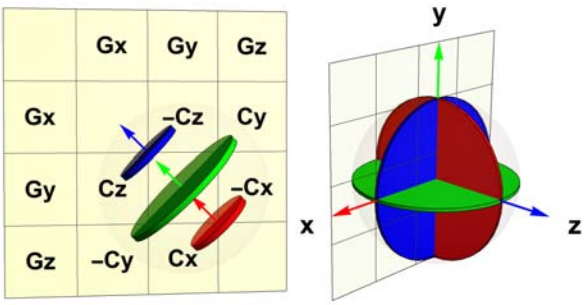



**Figure 1.** The circulating field, the C-field, can be labeled by the (row, col) component or by the orthogonal axis about which the (row, col) component circulates. For example, the (*x*,*z*) element is labeled $C_y$ and the (*z*, *x*) element is labeled $-C_y$ since both of these terms rotate about the y-axis; similarly for the other components. These rotations are shown abstractly in the representation of the field strength $F_{\mu\nu}$ matrix on the left. The right-hand illustration maps the three bivector diagrams into 3-space. Colors are used for visual convenience and for suggested correlation with $SU(3) \times SU(2) \times U(1)$ symmetry.

Isotopic gauge covariant field equations $f_{\mu\nu}$ are expressible in terms of Yang-Mills gauge field $\boldsymbol{b}_{\mu\nu}$

$$\boldsymbol{f}_{\mu\nu} = \underbrace{\frac{\partial \boldsymbol{b}_\mu}{\partial x_\nu} - \frac{\partial \boldsymbol{b}_\nu}{\partial x_\mu}}_{C_{\mu\nu}} - \underbrace{2m\boldsymbol{b}_\mu \times \boldsymbol{b}_\nu}_{mv_{\mu\nu}} \tag{17}$$

In this case the kinetic term of the Lagrangian, $\mathcal{L} = \frac{1}{4}\boldsymbol{f}_{\mu\nu} \cdot \boldsymbol{f}_{\mu\nu}$, will contain a product term proportional to $\left(C_{\mu\nu}\right)\left(mv_{\mu\nu}\right)$ and a quadratic term $\left(mv_{\mu\nu}\right)^2$. The scalar multiplier $\left(g/c^2\right)$ has dimension $l/m = \text{length/mass}$ hence $\left(g/c^2\right)\left(mv_{\mu\nu}\right)^2 \Rightarrow \left(mv^2\right)\cdot\text{length}$. The product term corresponds to $\boldsymbol{f}_{\mu\nu} \times 2m\boldsymbol{v}_{\mu\nu}$ which, in Equation (15), shows up as a new source term. In other words, our treatment of the gravitomagnetic gauge field matches Yang-Mills' original treatment.

*If it were obvious how to achieve mass gap at this point, it would have been solved in* 1954.

## 6. Higher-Order Self-Interaction

The Yang-Mills $\epsilon\left[A_\mu, A_\nu\right]$ term covers <u>all</u> gauge field component interactions, discussed above in terms of the original Yang-Mills paper. Yet neither mass gap nor quark confinement can be formulated successfully in this approach, so we examine a different self-interaction framework. The gravitomagnetic C-field has energy density, hence mass density, and circulates or rotates about an axis in space. The motion of the field, at any local point, results in momentum density at that point. But momentum density is the source current generating C-field circulation to begin with. Thus, the field itself induces more field and these fields interact; exactly what the Yang-Mills non-Abelian term is supposed to represent. *Therefore, we should investigate the real physical field interacting with itself instead of an abstract "internal" symmetry.* The mass density of the second order

induced circulation field is not equal to the mass density that induced the first circulation. The self-induced circulation is iterative; the first induced field induces a second order field circulation, and this, at any local point, induces a third order circulation, etc.

Physical spin is associated with circulation of the C-field; $\nabla\times\boldsymbol{C}$ represents bivector circulation, a spinning region of field such as a cross-section through a vortex, possessing angular momentum. Figure 2 illustrates first and second order induced fields caused by source momentum density, $p_0$. Higher order inductions of C-field circulations can be illustrated successively.

The first conclusion is that successive orders do *not* interact to any degree; the force $\boldsymbol{p}\times\boldsymbol{C}$ is always orthogonal to the velocity, hence the work done is zero: $\text{Work}=\int \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{x}=0$. Alternate orders, on the other hand, *do* interact, as they are parallel or anti-parallel. To schematically illustrate this, we take the tangent vectors to the circulation loops at the nearest and farthest points and "square the circle", using the straight lines as heuristic devices to facilitate the expression of forces involved via analogy with electromagnetic forces between parallel currents (Figure 3).

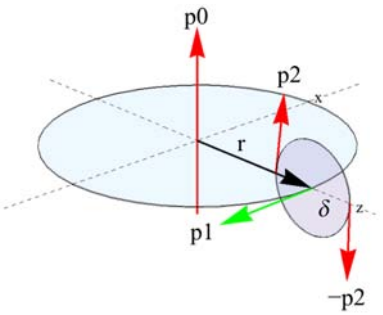



Figure 2. Momentum density $\boldsymbol{p}_0$ (red) induces C-field circulation at position $\boldsymbol{r}$. The C-field circulation at $\boldsymbol{r}$ yields momentum density $\boldsymbol{p}_1$ (green) orthogonal to $\boldsymbol{p}_0$. Momentum $\boldsymbol{p}_1$ induces the C-field at distance $\delta$ from $\boldsymbol{p}_1$. This induced C-field yields momentum density $\boldsymbol{p}_2$ (red) with components parallel and anti-parallel to $\boldsymbol{p}_0$.

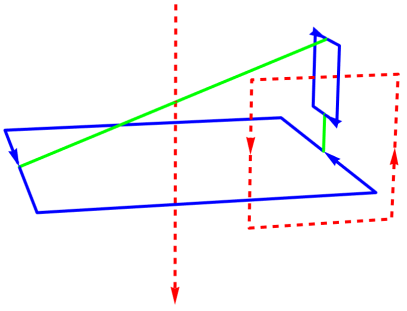


Figure 3. Focusing on (blue) loop1 and loop3 of the structure; source current and second order induction, loop2, are shown as red dashed lines. Since the loop3 bottom current is parallel to the rightmost current of loop1, the currents exert attractive forces upon each other, while top of loop3 is parallel to the current at the left of loop1 so the currents attract each other. The attractive force lines are shown in green. Similar same arguments apply to anti-parallel currents which exert repulsive forces (not shown).

The self-linking field formalism of Figure 2 shows that second-order induction reinforces the primary inducing agent, *i.e.*, local momentum density $\rho \boldsymbol{v}$. The electromagnetic force $\boldsymbol{F}_{ij}$ between two current elements $\mathrm{d}\boldsymbol{j}_i$ and $\mathrm{d}\boldsymbol{j}_j$ a distance $\boldsymbol{r}_{ij}$ apart guides us to write the gravitomagnetic equivalent.

$$\mathrm{d}\boldsymbol{F}_{ij} = \frac{\left[\mathrm{d}\boldsymbol{p}_j\right] \times \left[\mathrm{d}\boldsymbol{p}_i \times \boldsymbol{r}_{ij}\right]}{r_{ij}^3}. \tag{18}$$

Since $\mathrm{d}\boldsymbol{C}_i = \mathrm{d}\boldsymbol{p}_i \times \frac{\boldsymbol{r}_{ij}}{r_{ij}^3}$ where $\mathrm{d}\boldsymbol{p}_i$ is the mass current element inducing the field then $\mathrm{d}\boldsymbol{p}_j \times \mathrm{d}\boldsymbol{C}_i$ and Equation (18) is seen to be compatible with the Lorentz force law $\boldsymbol{F} = \boldsymbol{p} \times \boldsymbol{C}$ for the force on momentum $\boldsymbol{p}$ in gravitomagnetic field $\boldsymbol{C}$. In Figure 2, first-order C-field induction from momentum source density $\boldsymbol{p}_0$, is used to derive second order C-field induction from the momentum of the first-order field, $\boldsymbol{p}_1 \sim \boldsymbol{C}_1 \cdot \boldsymbol{C}_1$. Figure 3 focuses attention on loop1 and loop3 of the structure, showing the source current, and second order induction, loop2, as dashed lines. The bottom current in loop3 is parallel to the rightmost current of loop1, and therefore the currents exert attractive forces upon each other. Similarly, the current at the top of loop3 is parallel to the current at the left of loop1 and the two currents attract each other. The same arguments apply to the anti-parallel currents which exert repulsive forces. The above follows from

$$\mathrm{d}\boldsymbol{F}_{01} = \mathrm{d}\boldsymbol{p}_1 \times \mathrm{d}\boldsymbol{C}_0 = 0 \quad \text{since} \quad \boldsymbol{C}_0 \parallel \boldsymbol{p}_1 \tag{19}$$

$$\mathrm{d}F_{02} = \mathrm{d}\boldsymbol{p}_2 \times \mathrm{d}\boldsymbol{C}_0 \neq 0 \quad \text{since} \quad \boldsymbol{C}_0 \perp \boldsymbol{p}_2 \tag{20}$$

The force between $\boldsymbol{p}_0$ and $\boldsymbol{p}_1$ is zero since these mass density current flows are orthogonal to each other. On the other hand, the force acting between $\boldsymbol{p}_0$ and $\boldsymbol{p}_2$ is maximal or minimal according to whether these flows are parallel or anti-parallel.

This schematic organization guides calculation of the forces involved in the self-interaction of a turbulent primordial field. We seek first a qualitative understanding of dynamic behavior. All squares in the diagrams represent extensions of the tangent vectors depicted in Figure 3 and restore the dashed red loop2 to its true circular form. With this revision current loop3 should rotate about loop2, under the influence of the forces, eventually rotating into the *xy*-plane as depicted in Figure 4.

Loop3, shown in blue above loop1, is simply a slice through a torus surrounding loop2. It has no independent existence such that it can be pulled down into the plane. Nevertheless, if a "slice" is pulled into the plane, the field that *replaces* that slice will experience the same forces; the net result is a dynamic tension that tends to shrink the system of circulations into a lower energy configurational state. The final state of an arbitrary slice is depicted in Figure 5.

Despite having higher order constructions, the behavior is almost certainly governed by interactions between 1st and 3rd order induced circulations, as shown in Figure 4, consisting of the loop1 currents into and out of the page and

the two loop3 circulations, each with parallel currents into and out of the page. To formalize these interactions, we define the interaction between momentum density currents $p_i$ and $p_j$ as $f\left[p[i], p[j]\right]$ divided by the absolute distance between the currents and construct the interaction matrix over all six relevant currents, shown numbered in Figure 5.

## 7. Path Integrals over the Lattice

In Figure 3 a (blue) loop3 is vertically aligned over one leg of loop1. Figure 4 is a snapshot of loop3 rotating about loop2 from the initial vertical state ($\theta = 0$) to the horizontal state in the loop1 plane ($\theta = -\pi/2$). The loop is symmetric and supports an inverse image behavior from another loop3 on the left side of the diagram. To proceed from the initial state to the final state, we step through a sequence of rotations. The paths through the local space surrounding loop1 are traced out by rays originating on loop1 and rotating by $\mathrm{d}\theta_i$ rotations from $\theta = 0$ to $\theta = -\pi/2$. Two paths are traced—the lower leg of loop3, and the upper leg of loop3 as loop3 rotates from vertical to horizontal. This lattice of points defines the points at which we want to calculate forces between loops. Four snapshots of such lattice-based dynamics are shown in Figure 6.

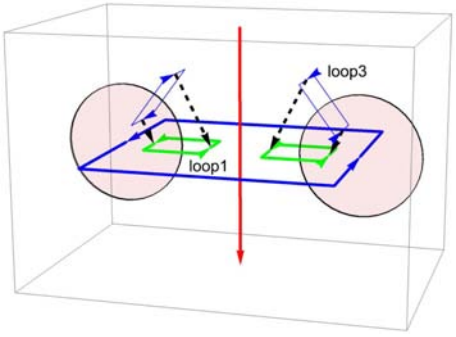



Figure 4. Cartoon snapshot depicting third-order loop (blue) dynamics interacting with first order loop (blue) of C-field circulation induced by (red) source momentum $\boldsymbol{p}_0$.

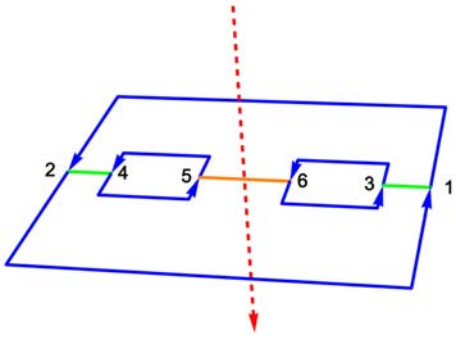



Figure 5. The result of dynamic forces acting on slices of higher order loops whose currents are numbered as shown. In this progression the configuration shown exerts attractive forces (green) between higher order loops and lower order loops, and repulsive forces (orange) between displaced higher order loops. This behavior follows at all orders.

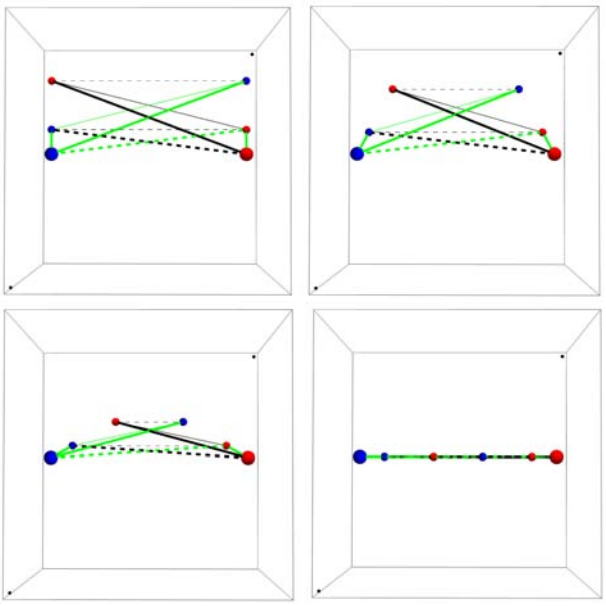


**Figure 6.** Shows positions and directions from initial vertical state represented by parameter $\theta = 0$, successively transforming to parameter $\theta = -\pi/2$. Attractive forces are shown as solid lines, with repulsive forces represented by dashed lines. Black lines represent forces to be calculated, while green lines represent symmetric forces; identical to forces corresponding to black lines.

A time-sliced adaptation of Figure 4 shows the relevant portions of the current loops directed into the page (red) and out of the page (blue). Attractive forces are shown as solid lines, repulsive forces by dashed lines. Black lines represent forces to be calculated, while green lines represent symmetric forces, identical to forces corresponding to black lines. Thick lines correspond to one power factor $\alpha$ while thin lines are interactions with factor $\alpha^2$. Figure 6 shows positions and directions from initial vertical state (parameter $\theta = 0$) successively transforming to $\theta = -\pi/2$.

Calculations of work done by the forces have the form $W_{ij}(\theta) = \boldsymbol{F}_{ij}(\theta) \cdot \mathrm{d}\boldsymbol{x}_i(\theta)$ where indices *i* and *j* vary from one to six as shown in Figure 5 and $\theta$ varies from 0 to −π/2 as currents 3 and 6 move from initial vertical position into the *xy*-plane. The displacement $\mathrm{d}\boldsymbol{x}_i(\theta) = \boldsymbol{x}_i(\theta) - \boldsymbol{x}_i(\theta + \mathrm{d}\theta)$. Examination shows that $|\mathrm{d}\boldsymbol{x}_5|$ and $|\mathrm{d}\boldsymbol{x}_6|$ are greater than $|\mathrm{d}\boldsymbol{x}_4|$ and $|\mathrm{d}\boldsymbol{x}_3|$ for the same $\mathrm{d}\theta$.

In Figure 6, for example, when currents 5 and 6 come together in the plane from initial vertical position, they oppose each other and the field between them increases, hence the energy density of the field increases, representing positive work shown by $W_{56}$. Currents 3 and 5, on the other hand, are parallel and attract each other, minimizing their joint field between them and reducing the energy, thus representing negative work, shown by $W_{35}$ in Figure 7.

The forces and displacements are calculated for every step of travel along the lattice path. The current force $\boldsymbol{F}_{ij}(\theta)$ applied over $\mathrm{d}\boldsymbol{x}_i(\theta)$ describes the work done for that step. The inner product $\boldsymbol{F}_{ij}(\theta) \cdot \mathrm{d}\boldsymbol{x}_i(\theta)$ is maximum when $\boldsymbol{F}_{ij}(\theta)$ and $\mathrm{d}\boldsymbol{x}_i(\theta)$ are parallel. Since we began calculations at $\theta = 0$, the ini-

tial $\mathrm{d}\boldsymbol{x}_1(0)$ is $(-\mathrm{d}x,0,0)$ while the initial force $\boldsymbol{F}_{ij}(0)=\alpha(-x,-y,0)$. The two vectors are not parallel. They become parallel when $\boldsymbol{F}_{ij}(\theta_k)\|\mathrm{d}\boldsymbol{x}_i(\theta_k)$. Thus, in Figure 8 the energy is seen to peak at $\theta_k\sim-\pi/6$. From that point onward, each successive step will lead to a lower energy state, and mass-energy density of the field structure becomes more "locked-in". In this way particles emerge with mass-energy greater than the vacuum state.

## 8. Brief Summary of Physics

The above physics is based on vorticity as the ubiquitous aspect of turbulent superfluid. Energy flows from large vortices to smaller vortices, which are circulating regions in the ultra-dense gravitomagnetic gauge field, with positive energy over a small region. The motion of the local field circulation induces further circulation and this in turn induces even higher order circulation. The topology is such that orders differing by one do not interact, whereas orders that differ by

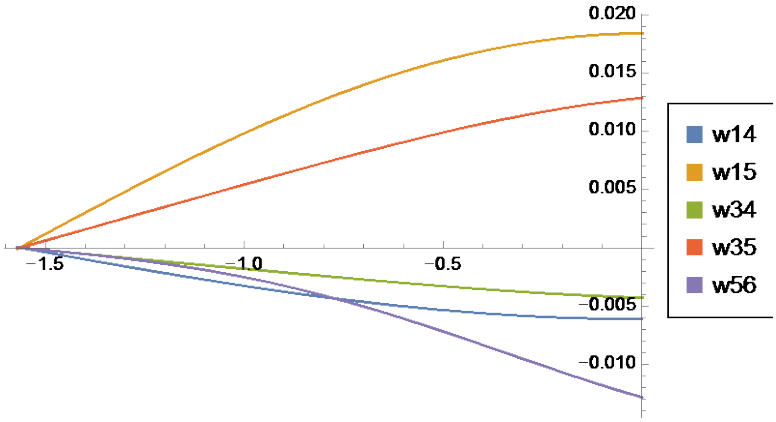



Figure 7. The work $W_{ij}(\theta)$ representing the interactive force $\boldsymbol{F}_{ij}(\theta)$ between momentum currents $i$ and $j$ at angles $\theta$. The horizontal axis runs from $\theta=0$ to $\theta=-\pi/2$ while the vertical axis represents work $W_{ij}(\theta)=\boldsymbol{F}_{ij}(\theta)\cdot\mathrm{d}\boldsymbol{x}_i(\theta)$. At $\theta=-\pi/2$ all displacements are down while all forces are horizontal, so no work is done at the final angle.

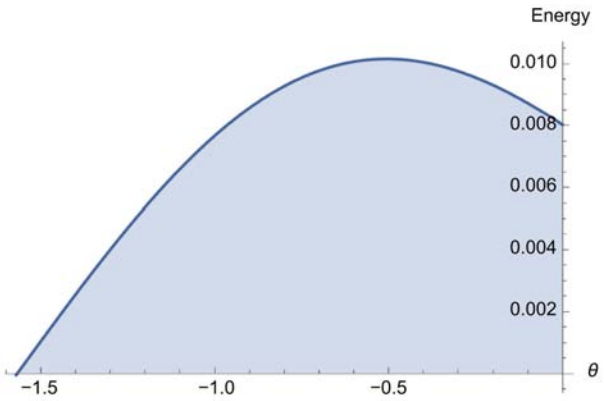



Figure 8. Summing all the (arbitrarily scaled) works involved, we find that the net energy decreases, thus the self-interaction of the field leads to a more stable configuration.

an even number *do* interact. We have employed a fractal structure and defined a path based on the relevant self-interactions. The forces act to move the induced flows into the primary circulation plane while shrinking the boundary of the field. This movement drives the system to a lower, but still positive, energy which is denser than the initial vortex energy. This ultra-dense stabilized spin energy is the "mass-gap" that yields particle mass greater than the vacuum state. We have thus shown that the primordial field "condenses" to a positive energy structure that we identify as a fundamental particle. We know that this particle will have quantized spin and quantized charge, and these aspects must be included to calculate the actual mass of the particle. This is in progress.

## 9. Discussion of Results

The years 1954 and 1964 witnessed revolutionary mathematics introduced into physics—Yang-Mills non-Abelian gauge theory and Hestenes' geometric calculus, supporting physical intuition relevant to Yang-Mills theory wherein the non-Abelian term $\left[A_\mu, A_\nu\right] \sim A_\mu \times A_\nu$ represents the interaction of the gauge field with itself. This term $\epsilon\left[A_\mu, A_\nu\right]$ is supposed to cover all the gauge field self-interactions, yet neither mass-gap nor confinement has been formulated successfully in this approach. "Free" gauge field propagation through space can conceivably self-interact, but nothing stable arises from such interactions. It is intuitively obvious that a "mass gap" can arise only locally, which seems to imply either a boundary or a local potential well. Alternatively, a local circulation may be invoked, almost certainly the reason for the focus on "angular momentum" in isospin space, which is abstract, based on approximate symmetry. As Weinberg states [18]: "*Many symmetries…were approximate because they weren't fundamental symmetries at all*; *they were just accidents*." Despite their approximate nature, there has been much focus on symmetry aspects in Yang-Mills. In "*Yang-Mills Origin of Gravitational Symmetries*" [19] many gravitational symmetries are derived linearly, including general covariance, two-form gauge invariance, local supersymmetry, and local chiral symmetry, following flat-space Yang-Mills theory. They remark that an important improvement would address the issue of dynamics as well as symmetry. Finally, "*we might speculate that the supergravity* $\phi_\mu$, *the left Yang-Mills* $V^i(L)$, *the right Yang-Mills* $A_\mu^{i'}(R)$ *and the spectator* $\Phi_{ii'}$ *live in different worlds with their own Lagrangians*." By contrast, our approach begins with one primordial world and nothing else, with an implicit self-interaction equation, the solution of which leads straightforwardly to higher order self-interactions. The stability of these interactions yields a mass gap that failed to appear from prior symmetry analysis.

In our derivation of Yang-Mills from the *Self-Interaction Principle*, the Heaviside C-field circulation is proportional to classical angular momentum. Equation (16) shows bivector-based $\epsilon\left[A_\mu, A_\nu\right] \Rightarrow m\left(A_\mu, A_\nu\right) = 2mv_{\mu\nu}$. The Yang-Mills non-Abelian term is proportional to angular momentum of the gravitational gauge field, rather than some isotopic "angular momentum" in abstract space;

our approach is based on a real physical field, with real angular momentum, formulated in terms of bivector rotations in real 3-space. $\epsilon\left[A_\mu, A_\nu\right]$ represents different interacting flows of the same field, locally distributed over space. The key Heaviside equation involving mass current density $\rho\boldsymbol{v}$ induces local circulation $\nabla\times\boldsymbol{C}=-\rho\boldsymbol{v}$ of the C-field with local energy density $\sim\boldsymbol{C}\cdot\boldsymbol{C}$ energy circulating with velocity $v'$ in the medium. This momentum density $\rho\boldsymbol{v}=(\boldsymbol{C}\cdot\boldsymbol{C})\boldsymbol{v}'$ will, in turn, induce second-order C-field circulation.

*This* is *Self-Interaction*; if stable, it will lead to a mass gap. Our approach demonstrates stability via iterative self-induction and shows that, while alternating inductions do not interact, even-order self-induced structures *do* interact, and do so in a manner that *increases the stability of the locally circulating field structure*. In this calculation the scale is unknown, and the coupling parameters are not rigorously specified, but the dynamical behavior of the model is correct. The forces between points on the lattice represent gauge field flows in and out of the plane of the paper which are time-linked. The dynamical energy exchange is a function of flow topology and distances. Global scale parameters will not change the direction of the energy evolution, only the magnitude of the effect. The resulting self-interactions lead to a lower energy and a greater density state. Like a skater pulling in her arms, a decreasing radius leads to increased angular velocity. For simplicity we have suppressed this "shrinkage" of the structure, but in reality we expect the radius of the circulating field to decrease and the local velocity $v'$ to increase, such that $m'v'r'$ is the conserved angular momentum, where $v'$ is the increased rotational velocity, $m'$ is the relativistic mass $m'=\gamma(v')m_0\sim\gamma(v')(\boldsymbol{C}\cdot\boldsymbol{C})$ and $r'$ is the reduced radius. Quantized angular momentum ($\sim\hbar$) should prevent this circulation from shrinking to an infinitely dense "point" particle, and therefore should evolve to a finite sized toroidal field structure whose mass spectrum is based on parameters to be specified, but whose existence as a stable field structure has been demonstrated. Following papers will address half-integer spin and electric charge aspects of the particle, however they will inevitably trace back to this re-interpretation of the Yang-Mills term representing non-Abelian self-interaction: $\epsilon\left[A_\mu^{(i)}, A_\nu^{(i+2)}\right]$ where ($i$) refers to induction order.

## 10. Summary

Our goal has been to formulate higher-order self-interaction of the gauge field and re-interpret the non-Abelian term based on this, to derive fermions from the gravitational gauge field. Previous papers have shown the derivation of Heaviside's equations from an exact principle, the *Self-Interaction Principle*, is equivalent to Einstein's nonlinear field equations derived from the *Equivalence Principle*, and have treated general relativity-based problems such as *Quasi-Local Mass*. Key is that the Heaviside derivation is field-strength independent, whereas Einstein's derivation erroneously implies that Heaviside is a "weak field approximation".

A *mass-based* understanding of gravity, as well as the *weak field approximation* misunderstanding, causes physicists to generally ignore gravity in particle physics. A *mass-density-based* understanding of gravity leads to a gravitational basis for particle physics. Burinskii [20] has suggested that particles arise from gravity with the structure of the Kerr black holes, while Christian and Diether [21] suggest particle radii on the order of the Planck length. In other words, mass densities associated with the big bang are effectively limitless. Huang [22], Volovik [23], and others view the primordial field as a *superfluid*. Circa 2006 physicists at the LHC were expecting a *quark gas* from heavy-ion collisions but instead [24]

"*It is well known that the properties of the Yang-Mills plasma turned out to be unexpected…the plasma is similar rather to an ideal liquid than to a gluon gas interacting perturbatively*."

They conclude with an analogy between phenomena in Yang-Mills theory with physics of *superfluidity*. Our underlying premise has been the *superfluid nature of the primordial field*, with ultra-dense fields, in which we identify higher-order self-induction modes.

Einstein (1919) asked "*Do gravitational fields play an essential part in the structure of the elementary particles of matter*?", suggesting the possibility of a theoretical construction of matter out of gravitational field and electromagnetic field alone. From't Hooft's perspective: "*Einstein's theory of general relativity has a mathematical structure very similar to Yang-Mills theory*." And Zee remarks: "*there is increasing evidence that the Einstein theory of gravity is just Yang-Mills squared*." Yet the *Millennium prize* declares that:

"*Yang-Mills theory is now the foundation of most of elementary particle theory, but the mathematical foundation is still unclear*."

*Physical* ideas may have been the source of Yang-Mills failure on key issues, not *mathematical* ideas. *Density*-based gravity may open realms of physics to gravitational phenomena that have been overlooked since Newton. This paper has presented a density-based re-interpretation of Yang-Mills gauge field *self-interaction* leading to stable gravitational gauge field structures to explain the mass gap. Future papers will explore half-integer spin and genesis of electric charge. These two issues should allow derivation of mass of fermions whose mass-gap was derived herein.

## Conflicts of Interest

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

## References

[1] Zee, A. (2003) Quantum Field Theory in a Nutshell. Princeton University Press, Princeton.

[2] Feynman, R. (1995) Feynman Lectures on Gravitation. Westview Press, Boulder.

[3] Klingman, E. (2020) *Journal of Modern Physics*, **12**, 65-81. https://doi.org/10.4236/jmp.2021.122007

[4] Almeida, J. (2008) Ether Spacetime and Cosmology Vol. III. Apeiron Pub., Montreal, 257.

[5] Hestenes, D. and Sobcyzk, G. (1984) Clifford Algebra to Geometric Calculus. Reidel Publishing Company, Dordrecht, 242. https://doi.org/10.1007/978-94-009-6292-7

[6] Heaviside, O. (1893) *The Electrician*, **31**, 81-82.

[7] Klingman, E. (2022) *Journal of Applied Mathematics and Physics*, **10**.

[8] O'Raifeartaigh, L. (1997) The Dawning of Gauge Theory. Princeton Univ Press, Princeton. https://doi.org/10.1515/9780691215112

[9] Klingman, E. (2022) *Journal of Modern Physics*, **13**, 368-384. https://doi.org/10.4236/jmp.2022.134026

[10] Huang, K. (2007) Fundamental Forces…the Story of Gauge Fields. World Scientific, Hackensack. https://doi.org/10.1142/6447

[11] Penrose, R (2021) Conversations on Quantum Gravity. Cambridge University Press, Cambridge.

[12] Klingman, E. (2021) *Journal of Modern Physics*, **12**, 1190-1209. https://doi.org/10.4236/jmp.2021.129073

[13] Klingman, E. (2022) *Journal of Modern Physics*, **13**, 347-367. https://doi.org/10.4236/jmp.2022.134025

[14] Klingman, E. (2021) *Journal of High Energy Physics, Gravitation and Cosmology,* **7**, 936-948. https://doi.org/10.4236/jhepgc.2021.73054

[15] Klingman, E. (2020) *Journal of Modern Physics*, **11**, 1950-1968. https://doi.org/10.4236/jmp.2020.1112123

[16] Yang, C. and Mills, R. (1954) *Physical Review Letters,* **95**, 631.

[17] Einstein, A. and deHaas, W. (1915) *KNAW Proceedings*, **18**, 696-711.

[18] Weinberg, S. (2005) The Making of the Standard Model. In: 't Hooft, G., Ed., 50 *Years of Yang-Mills Theory*, World Scientific, Singapore, 99-117. https://doi.org/10.1142/5601

[19] Anastasiou, A. (2014) *Physical Review Letters,* **113**, Article ID: 231606. https://doi.org/10.1103/PhysRevLett.113.231606

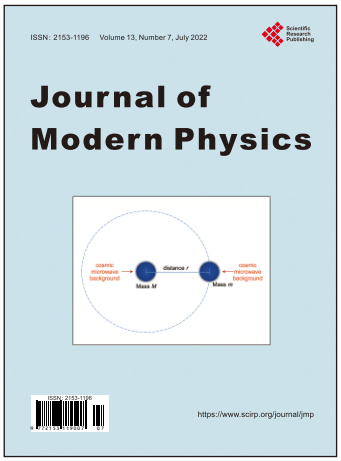
[20] Burinskii, A. (2017) Weakness of Gravity as Illusion….

[21] Christian, J. and Diether, F. (2019) Evidence of Matter as a Proof of the Existence of Gravitational Torsion.

[22] Huang, K. (2017) A Superfluid Universe. World Scientific, Singapore. https://doi.org/10.1142/10249

[23] Volovik, G. (2009) The Universe in a Helium Droplet. Oxford Science Pub., Oxford. https://doi.org/10.1093/acprof:oso/9780199564842.001.0001

[24] Chernodub, M. and Zakharov, V. (2007) *Physical Review Letters,* **98**, Article ID: 082002. https://doi.org/10.1103/PhysRevLett.98.082002




# 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**