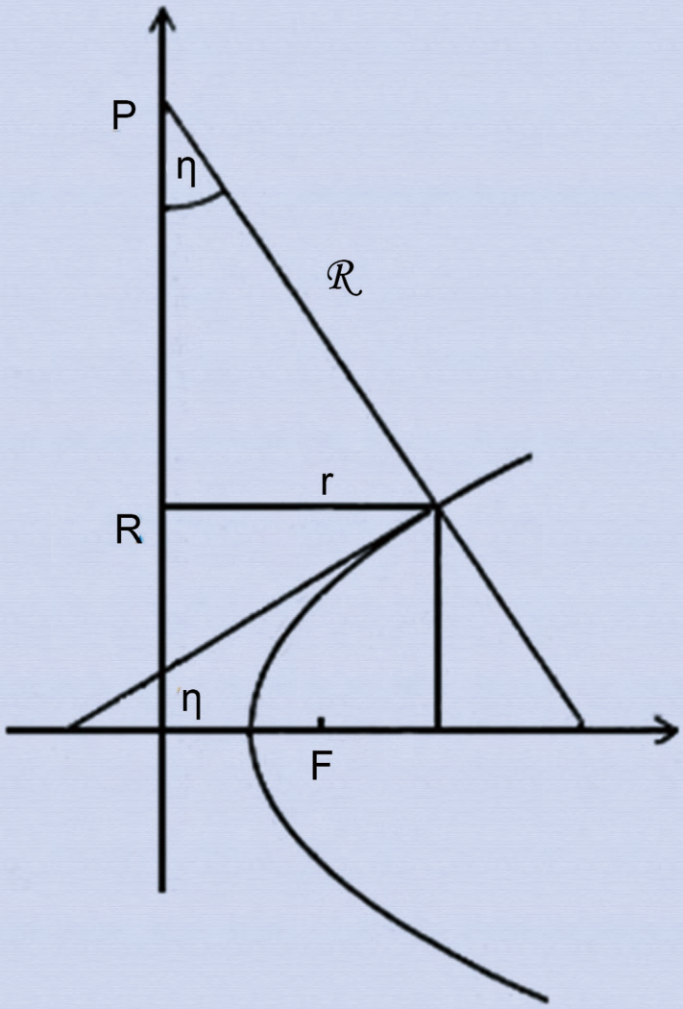
ISSN: 2153-1196 Volume 11, Number 3, March 2020

Scientific Research Publishing

Journal of Modern Physics



ISSN: 2153-1196
9 772153 119007 03

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

Journal of Modern Physics, 2020, 11, 343-485
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Table of Contents

Volume 11 Number 3 March 2020

The figure on the front cover is from the article published in Journal of Modern Physics, 2020, Vol. 11, No. 3, pp. 355-364 by Rainer Burghardt.

Journal of Modern Physics (JMP)
Journal Information

SUBSCRIPTIONS

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

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

SERVICES

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

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

COPYRIGHT

PRODUCTION INFORMATION

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

Journal of Modern Physics, 2020, 11, 343-354
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Entropy Production in a Non-Isolated Thermodynamic System Taking into Account Regular Factors of Nonrandom Nature

A. Yu. Khlestkov, Yu. A. Khlestkov, N. Yu. Lukashina, M. Yu. Lukashin, P. Yu. Lukashin

National Research Nuclear University MEPhI, Moscow, Russia
Email: khlestkov@yandex.ru

How to cite this paper: Khlestkov, A.Y., Khlestkov, Y.A., Lukashina, N.Y., Lukashin, M.Y. and Lukashin, P.Y. (2020) Entropy Production in a Non-Isolated Thermodynamic System Taking into Account Regular Factors of Nonrandom Nature. *Journal of Modern Physics*, **11**, 343-354.
https://doi.org/10.4236/jmp.2020.113021

Received: January 23, 2020
Accepted: February 28, 2020
Published: March 2, 2020

Abstract

The work illustrates the impossibility of decreasing entropy in a strictly random thermodynamic process in a non-isolated system using the example of heating a planet by solar radiation flux without and taking into account its rotation around its own axis. That is, the second law of thermodynamics formulated for isolated systems continues to govern such systems. We have shown that in order to achieve a stationary state at lower values of temperature and entropy far from thermodynamic equilibrium at a maximum of temperature and entropy, it is necessary to have regular factors of nonrandom nature, one of which in this example is the rotation of the planet around its own axis. This means that the reason for the appearance of ordered structured objects in non-isolated thermodynamic systems is not the random process itself, but the action of dynamic control mechanisms, such as periodic external influences, nonlinear elements with positive feedback, catalysts for chemical reactions, etc. We present the plots with dependences of temperature and entropy versus time in non-isolated systems with purely random processes and in the presence of a control factor of non-random nature-rotation.

Keywords

Random Process, Non-Isolated Systems, Entropy, Ordered Structures, Regular Factors of Non-Random Nature

1. Introduction

When studying the problem of origin of ordered structures: elementary particles, nuclei of atoms, crystals, biomolecules, cells, ..., basically one model is consi-

dered: their occurrence in a random process [1] [2] [3]. With this "self-organization" the disorder should decrease. This means that entropy as a measure of disorder in a thermodynamic process should decrease. But the second law of thermodynamics prohibits a decrease in entropy in random processes in isolated systems. Therefore, our certain hopes are associated with the openness of real systems allowing the presence of ingoing and outgoing flows of heat (closed systems), heat and particles (open systems) that do not fall under action of the second law. It is believed that there it is possible, in principle, to reduce entropy due to its "removal" by the outgoing flows, and, therefore, the creation of conditions for the appearance of ordered structures in a random process.

We consider the well-known "entropy pump" model which describes the process of decreasing entropy in the Earth's ecosphere. In this model, a change in the entropy S with a speed $\mathrm{d}S/\mathrm{d}t$ occurs in a non-isolated system in the course of a stationary process, when the ingoing ordered low-entropy heat flow P_s at solar temperature T_s is compared in equilibrium with the outgoing flow $P = P_s$. The latter is already more randomized (highly entropic) and having a lower temperature $T < T_s$. In this case, the rate of change in the entropy of the system becomes negative [4]:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = P_s\left(\frac{1}{T_s} - \frac{1}{T}\right) < 0 \tag{1}$$

There are two comments regarding formula (1). Firstly, concerning a stationary process in a homogeneous system, which in this example is a planet (neglecting the atmosphere) heated by the solar radiation flux P_s, when two temperatures arise, T_s and $T < T_s$. Since temperature is a function of the state, then in an equilibrium thermodynamic state the system can have only one temperature equal to the temperature of the ingoing heat flow T_s.

Secondly, the entropy S itself is also a function of the state, therefore, in equilibrium state it must also be constant, that is, its rate of change $\mathrm{d}S/\mathrm{d}t$ must be zero.

Let's see how we can obtain formula (1), and what it means in the classical thermodynamics of quasistatic systems, *i.e.* those in which a local thermodynamic equilibrium can be established in a non-equilibrium process in a small neighborhood of each point.

The simplest model was used (**Figure 1**), which we will describe in detail below. As a result of solving her system of differential equations, the following results were obtained:

- Taken with the opposite sign, formula (1) describes in fact the usually neglected production of entropy at the input to the system at the boundary of a body and the environment during thermalization of the ingoing heat flow P_s and its randomization during interaction between the photons belonging to the ingoing solar radiation and atoms of the crystalline lattice at the surface of the system (the planet's soil, in this example);

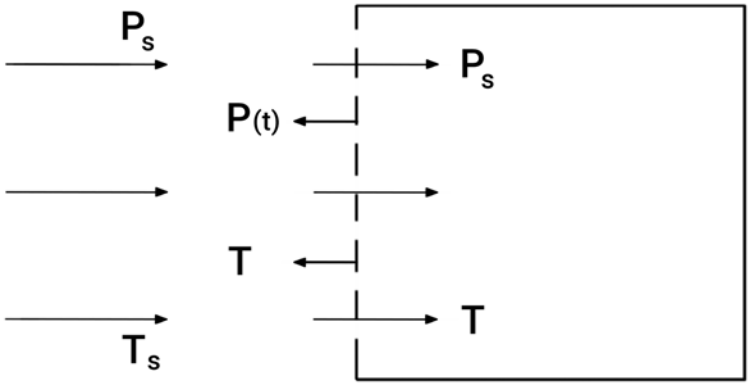


Figure 1. A non-isolated thermodynamic system with an input point outside the body. Thermalization of the ingoing heat flow at the system boundary.

- Taking into account this production of entropy, despite its removal by the outgoing heat flow during the heating of the planet, the entropy of this non-isolated system does not decrease. That is, the second law of thermodynamics remains valid for non-isolated homogeneous systems, if all the processes occurring in it are random;
- At $t \to \infty, T \to T_s, \frac{\mathrm{d}S}{\mathrm{d}t} \to 0, S(t) \to S_{\infty \max}$;

that is, a new asymptotic equilibrium state of the system is achieved at the maximum temperature $T = T_s$ and the entropy $S = S_{\infty \max}$. There is no entropy pump here.

But a paradox arises: in fact, however, real planets without a significant atmosphere (such as Mercury, Mars, and Pluto) really have two temperatures: T_s is the total radiation temperature of ingoing solar radiation and $T_\infty < T_s$ is the stationary temperature of the planet surface layer heated by solar heat. Therefore, they are not described by this thermodynamic model.

It is clear, what is the reason for this difference. In the considered illustrative example, the influence of the planet's rotation around its axis on the temperature and entropy balance is not taken into account. Therefore, we have a periodic change of day and night, at which heating, simultaneous emission of heat by the body and an increase in entropy occur on the day side of the planet, while on its night side the planet only radiates heat without generating entropy.

This leads to the fact that the stationary state is achieved by a non-isolated system at lower values of temperature and entropy, far from the thermodynamic maximum.

We took into account this rotation (in the general case, a periodic pulsed heat flow from the heater (**Figure 2**)) and we have obtained the expected result: rotation of the planet around its axis decreases the stationary temperature of its surface and entropy.

The calculation of the day and night asymptotic temperature values $T_{d\infty}$ and $T_{n\infty}$ for the planets mentioned above showed good agreement with the experimental data (see **Table 1**).

Table 1. Calculation of the Day and Night temperatures for the three planets.

Planet	Mercury	Mars	Pluto
Distance from the Sun, million km (a.u.)	57 (0.39)	230 (1.5)	6000 (40)
Solar constant, W/m^2	4200	590	1
Rotation period, Earth days	58.6	1.03	6.39
Total radiation temperature due to solar radiation, K	630	320	63
Daily experimental temperature $T_{d\infty}$, K	500 - 700	200 - 300	40
Daily calculated temperature $T_{d\infty}$, K	630	270	50
Night experimental temperature $T_{n\infty}$, K	100	150 - 250	-
Night calculated temperature $T_{n\infty}$, K	150	260	50

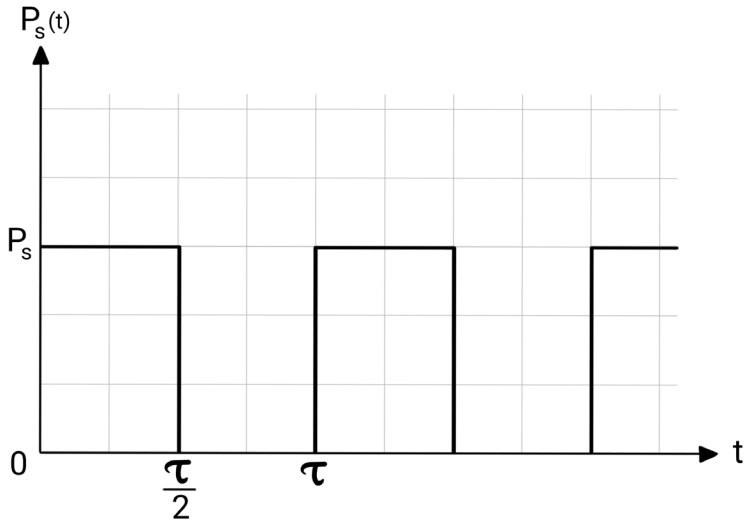


Figure 2. The periodic heat flow $P_s(t)$ as the regular control factor.

Thus, in addition to the conclusion that entropy still does not decrease in a non-isolated system with purely random processes, which means that no self-organization phenomena are possible in purely random processes, one more result is obtained: the entropy of a stationary non-isolated system can become less than a thermodynamic maximum only under the influence of regular factors of a non-random dynamic nature (in this example, pulsed semi-periodic heat transfer, starting from $t = 0$).

Hence the following conclusion is possible: the reason for the decrease in disorder and the appearance of structured objects is not the random process itself, but the action of not random ordering mechanisms, moreover of external origin: periodic external influences, such as nuclear or chemical reactions, nonlinear dynamic mechanisms with positive feedback, etc.

This result is consistent with the Prigogine hypothesis about the existence of control parameters [1] [3]. We only note according to the result, that these parameters can be of a regular nature, and their nature needs to be investigated.

For example, the considered mechanism of periodic interruption of the in-

going heat flow is provided by the rotation of the planets during the gravitational interaction of bodies in the universe.

2. Thermodynamic Model without Rotation

In Figure 1, a low-entropic heat flow P_s with a total radiation temperature T_s falls on the boundary of a non-isolated system and becomes randomized when interacting with the surface of a body (planet). Other dissipative processes (thermal conductivity, diffusion, viscosity ...) are neglected.

This heat flow begins to heat the body from the initial temperature T_0 to $T(t)$. When heated, the body begins to radiate a return heat flow into an external environment with increasing power $P(t)$ until a new equilibrium state with the entropy $S(T)$ is reached, which, like the body temperature $T(t)$, will be calculated from the following balance of entropy [5] [6]:

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= \frac{\mathrm{d}S_{ext}}{\mathrm{d}t} + \frac{\mathrm{d}S_{int}}{\mathrm{d}t} = \left(\frac{\mathrm{d}S_{inp}}{\mathrm{d}t} - \frac{\mathrm{d}S_{out}}{\mathrm{d}t}\right) + \frac{\mathrm{d}S_{int}}{\mathrm{d}t} \\ &= \left(\frac{P_s}{T_s} - \frac{P(t)}{T}\right) + \left(\frac{P_s}{T} - \frac{P_s}{T_s}\right) = \frac{1}{T}\frac{\mathrm{d}U + \delta A}{\mathrm{d}t} \\ &= \frac{1}{T}c_v M\frac{\mathrm{d}T}{\mathrm{d}t} + \frac{1}{T}\frac{\delta A}{\mathrm{d}t}\end{aligned} \tag{2}$$

Here S_{ext} is the entropy transferred by heat flows: ingoing, S_{inp}, with speed $\frac{P_s}{T_s}$, and outgoing, S_{out}, with speed $\frac{P(T)}{T}$; S_{int} is the entropy produced inside the system during thermalization of the ingoing heat flow at a speed equal to the increment $\left(\frac{P_s}{T} - \frac{P_s}{T_s}\right)$; $\mathrm{d}U$ is the differential of internal energy, c_v is the heat capacity of a body with a heated mass M; δA is a work performed by the body in the process of thermalization of the ingoing heat flow.

For simplicity, we assume that all heat flows are radiated by a completely black body:

$$P_s = \Pi\sigma T_s^4, \tag{3}$$

$$P = \Pi\sigma T^4, \tag{4}$$

where Π -surface area of the heated body, σ-Stefan-Boltzmann constant.

From relations (2)-(4), it follows:

$$\frac{\mathrm{d}S_{ext}}{\mathrm{d}t} = \Pi\sigma\left(T_s^3 - T^3\right) = c_v M\frac{1}{T}\frac{\mathrm{d}T}{\mathrm{d}t}, \tag{5}$$

$$\frac{\mathrm{d}S_{int}}{\mathrm{d}t} = P_s\left(\frac{1}{T} - \frac{1}{T_s}\right) = \frac{1}{T}\frac{\delta A}{\mathrm{d}t}, \tag{6}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\Pi\sigma}{T}\left(T_s^4 - T^4\right). \tag{7}$$

From (5) we obtain the equation for $T(t)$:

$$\frac{\Pi\sigma}{c_v M}\mathrm{d}t = \frac{1}{T}\frac{\mathrm{d}T}{T_s^3 - T^3} \tag{8}$$

and from (7), substituting into it the time differential $\mathrm{d}t$ from (8), we obtain the equation for $S(T)$:

$$\mathrm{d}S = c_v M \frac{T_s^4 - T^4}{T^2\left(T_s^3 - T^3\right)}\mathrm{d}T. \tag{9}$$

For convenience, we introduce dimensionless quantities (with a tilde)

$$\begin{aligned} \mathrm{d}t &= \frac{\tau}{2}\mathrm{d}\tilde{t}, \\ \mathrm{d}T &= T_s \mathrm{d}\tilde{T}, \\ \mathrm{d}S &= c_v M \mathrm{d}\tilde{S}. \end{aligned} \tag{10}$$

Substituting (10) into (8) and (9), we finally obtain the differential equations for $\tilde{T}(\tilde{t})$ and $\tilde{S}(\tilde{T})$ in the dimensionless form:

$$\alpha \mathrm{d}\tilde{t} = \frac{1}{\tilde{T}}\frac{\mathrm{d}\tilde{T}}{1-\tilde{T}^3}, \tag{11}$$

$$\mathrm{d}\tilde{S} = \frac{1}{\tilde{T}^2}\frac{1-\tilde{T}^4}{1-\tilde{T}^3}\mathrm{d}\tilde{T}, \tag{12}$$

where the dimensionless constant is

$$\alpha = \frac{\Pi\sigma\tau}{2c_v M}T_s^3. \tag{13}$$

3. Non-Isolated System without Rotation

Integrating Equation (11), Equation (12) with the initial conditions:

$$\tilde{t} = \tilde{t}_0, \tilde{T} = \tilde{T}_0, \tilde{S} = \tilde{S}_0(t_0) \tag{14}$$

and omitting hereinafter all tildes for dimensionless temperatures $T(t)$ and entropy increment:

$$\Delta S(t) = S(t) - S_0(t_0) \tag{15}$$

we get the following solutions:

$$T(t) = \left(1 - \left(T_0^{-3} - 1\right)\mathrm{e}^{-3\alpha(t-t_0)}\right)^{-\frac{1}{3}}, \tag{16}$$

$$\Delta S(T) = \frac{1}{2}\left(\ln\frac{1+T+T^2}{1+T_0+T_0^2} + 2\left(\frac{1}{T_0} - \frac{1}{T}\right) - \frac{2}{\sqrt{3}}\left(\operatorname{arctg}\frac{2T+1}{\sqrt{3}} - \operatorname{arctg}\frac{2T_0+1}{\sqrt{3}}\right)\right). \tag{17}$$

Equation (11), Equation (13) and their solutions (16), (17) are valid within the framework of the classical nonrelativistic thermodynamics for any non-isolated homogeneous systems in which thermodynamic random processes occur. But in the capacity of an illustrative example explaining the meaning of the quantities introduced in the model, we consider heating a planet of mass M with specific heat capacity of the heated layer c_v with surface area Π by solar radiation with a heat flow P_s and with a total radiation temperature at the input to the system T_s, which is then reradiated by the planet into the surrounding vacuum by the heat

flow $P(T)$ at a changing temperature $T(t)$ until a new equilibrium state is reached at $t \to \infty$.

From the solutions (16), (17) it directly follows that both the temperature $T(t)$ and the entropy $S(T(t))$ are monotonically increasing functions which depend on time t, reaching their equilibrium values T_∞ and S_∞ in the asymptotics $t \to \infty$ corresponding to the thermodynamic maximum (Figure 3, Figure 4, curves at $\omega = 0$).

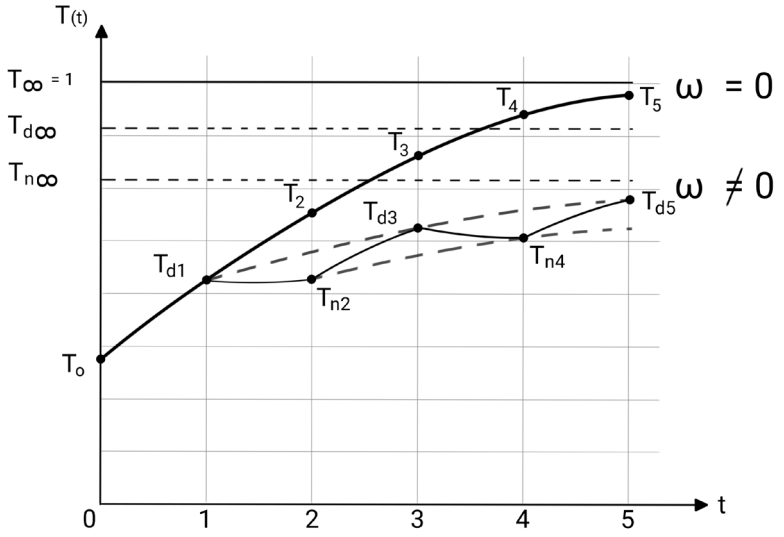


Figure 3. The dependence of the dimensionless body (planet) temperature $T(t, T_0)$ of a non-isolated thermodynamic system in the absence of rotation ($\omega = 0$) and in the presence of rotation ($\omega \neq 0$) (periodic heat flow at the input) on the day and night sides.

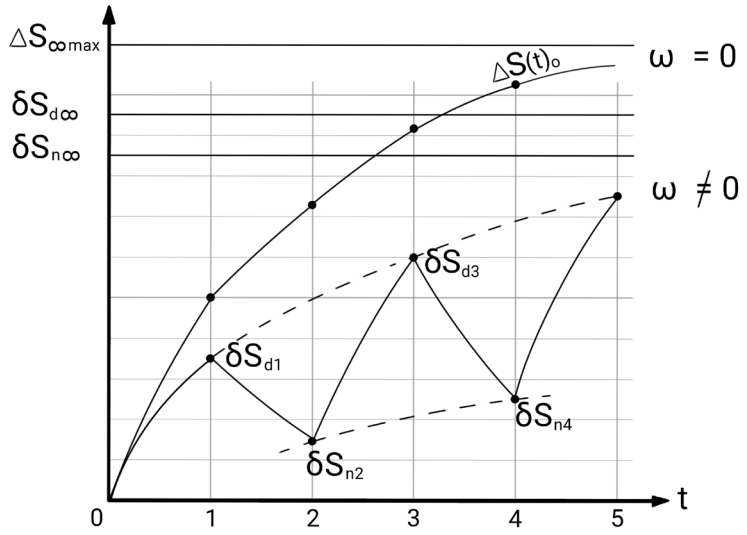


Figure 4. The entropy increment ΔS of a non-isolated system in a random process ($u = 0$, there is no rotation as a control factor) and increments of entropy on the day and night sides of the planet relative to its initial value $S_0(T_0)$ in the presence of a control factor of non-random nature (planet rotation around its axis, $\omega \neq 0$). $\Delta S_{\infty\,\max}$ is the dimensionless thermodynamic maximum of the entropy increment in a random process without controlling factors of non-random nature; $\delta S_{d\infty}$ and $\delta S_{n\infty}$ are the asymptotic stationary increments of entropy on the day and night sides of the planet, taking into account its rotation.

Thus, the second law of thermodynamics continues to govern at least for the considered class of non-isolated systems and random processes in them. This situation is shown in the graphs $T(t)$ and $\Delta S(t)$ in Figure 3, Figure 4. The asymptotic values of temperature and entropy are shown at $t \to \infty$, in a state of new thermodynamic equilibrium:

$$\begin{aligned} T(t) &\to T_\infty = 1, \\ \Delta S(T) \to \Delta S_\infty(1) &= \frac{1}{2}\left(\ln 3 - \ln\left(T_0^2 + T_0 + 1\right) + 2\left(\frac{1}{T_0} - 1\right) \right. \\ &\left. - \frac{2}{\sqrt{3}}\left(\frac{\pi}{3} - \operatorname{arctg}\frac{1+2T_0}{\sqrt{3}}\right)\right) > 0. \end{aligned} \tag{18}$$

4. Accounting for Body Rotation Being a Regular Factor of Non-Random Nature

Before you begin to format your It is possible to lower the entropy of the body to values less than its thermodynamic maximum by introducing a regular factor into the system, for example, periodically interrupting the ingoing heat flow $P_s(t) = P_s(t+\tau)$ in the considered example, taking into account the rotation of the planet around its axis (Figure 2, τ is the planet's day length (day + night)).

5. The Night Side of the Planet

On the day side, Equation (11), Equation (12) and their solutions (16), (17) remain valid. There is no ingoing heat flow on the night side, therefore,

$$P_s = 0, \frac{\mathrm{d}S_{inp}}{\mathrm{d}t} = 0, \frac{\mathrm{d}S_{int}}{\mathrm{d}t} = 0, \frac{\delta A}{\mathrm{d}t} = 0, \tag{19}$$

whence it follows from (2) (dimensional record):

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d}S_{ext}}{\mathrm{d}t} = -\frac{\mathrm{d}S_{out}}{\mathrm{d}t} = -\frac{P(T)}{T} = -\Pi\sigma T^3 = \frac{1}{T}c_v M\frac{\mathrm{d}T}{\mathrm{d}t}. \tag{20}$$

The dimensionless equations follow from (20):

$$-\alpha \mathrm{d}t = \frac{\mathrm{d}T}{T^4}, \tag{21}$$

$$\mathrm{d}S = \frac{\mathrm{d}T}{T} \tag{22}$$

And their solutions:

$$T(t) = \left(3\alpha(t - t_*) + T_*^{-3}\right)^{-\frac{1}{3}}, \tag{23}$$

$$\Delta S = S(T) - S_*(T_*) = \ln\frac{T}{T_*}, \tag{24}$$

where T_* and S_* are the temperature and entropy at the beginning of the night at $t = t_*$.

6. The Decrease in Entropy When a Periodic Factor of Non-Random Nature Is in Effect

Consider the temperatures T_d and T_n and the entropies S_d and S_n at the ends of the day and night sections (Figure 3, Figure 4, lower curves, the initial value T_0 is given):

Day: $2n \le t \le 2n+1,\ n = 0,1,2,\cdots$

$$T_{d\,2n+1} = \left(1 + \left(T_{n\,2n}^{-3} - 1\right)\mathrm{e}^{-3\alpha}\right)^{-\frac{1}{3}}, \tag{25}$$

$$\begin{aligned}\Delta S_{d\,2n+1} &= S_{d\,2n+1} - S_{n\,2n} \\ &= \frac{1}{2}\left(\ln\frac{1+T_{d\,2n+1}+T_{d\,2n+1}^2}{1+T_{n\,2n}+T_{n\,2n}^2} + 2\left(\frac{1}{T_{n\,2n}} - \frac{1}{T_{d\,2n+1}}\right)\right. \\ &\quad \left. - \frac{2}{\sqrt{3}}\left(\operatorname{arctg}\frac{2T_{d\,2n+1}+1}{\sqrt{3}} - \operatorname{arctg}\frac{2T_{n\,2n}+1}{\sqrt{3}}\right)\right).\end{aligned} \tag{26}$$

Night: $2n+1 \le t \le 2n+2$

$$T_{n\,2n+2} = \left(3\alpha + T_{d\,2n+1}^{-3}\right)^{-\frac{1}{3}}, \tag{27}$$

$$\Delta S_{n\,2n+2} = S_{n\,2n+2} - S_{d\,2n+1} = \ln\frac{T_{n\,2n+2}}{T_{d\,2n+1}}. \tag{28}$$

In formulas (26)-(28), the first n means "night", and the second $n \in 0,1,2,\cdots$ means integers.

$$\delta S_{d\,2n+1} = S_{d\,2n+1} - S_0 = S_{n\,2n} + \Delta S_{d\,2n+1} - S_0;$$

$$\delta S_{n\,2n+2} = S_{n\,2n+2} - S_0 = S_{d\,2n+1} + \Delta S_{n\,2n+2} - S_0.$$

Their dependence on time is shown in Figure 4.

At $n \to \infty$, a rotating planet changes to its new stationary state at constant day and night asymptotic temperatures $T_{d\infty}$, $T_{n\infty}$ and entropies $\delta S_{d\infty}$, $\delta S_{n\infty}$:

$$T_{d\infty} = \left(1 + \frac{3\alpha\mathrm{e}^{-3\alpha}}{1-\mathrm{e}^{-3\alpha}}\right)^{-\frac{1}{3}}, \tag{29}$$

$$T_{n\infty} = \left(1 + \frac{3\alpha}{1-\mathrm{e}^{-3\alpha}}\right)^{-\frac{1}{3}}, \tag{30}$$

$$\begin{aligned}\Delta S_{d\infty} &= \frac{1}{2}\left(\ln\frac{1+T_{d\infty}+T_{d\infty}^2}{1+T_{n\infty}+T_{n\infty}^2} + 2\left(\frac{1}{T_{n\infty}} - \frac{1}{T_{d\infty}}\right)\right. \\ &\quad \left. - \frac{2}{\sqrt{3}}\left(\operatorname{arctg}\frac{2T_{d\infty}+1}{\sqrt{3}} - \operatorname{arctg}\frac{2T_{n\infty}+1}{\sqrt{3}}\right)\right) \\ &> 0,\end{aligned} \tag{31}$$

$$\Delta S_{n\infty} = \ln\frac{T_{n\infty}}{T_{d\infty}} < 0. \tag{32}$$

Their graphs are presented in Figure 3, Figure 4. Two important results follow from these data:

- The increments of temperature and entropy on the day side are always greater than the decreases in temperature and entropy on the night side; the increment of entropy for a period, that is, for one revolution, always grows;
- At $t \to \infty$ ($n \to \infty$), the stationary asymptotic temperature and entropy increments, $T_{d\infty}$, $T_{n\infty}$ and $\delta S_{d\infty}$, $\delta S_{n\infty}$ are less than their maximum equilibrium values in a purely random process.

In principle, if to extend the night state in time (pause when the body heats up) and reduce the daytime state in time, we can achieve a decrease in temperature, and therefore, entropy, according to the third law of thermodynamics (Nernst theorem [5]), down to zero. But we will not do this in this work, this is a separate task that requires complicating the model by taking into account the dependence of the heat capacity of the body on temperature and fulfilling the condition: at $t \to 0$, $\lim c_v(T) = 0$. The results will not be affected by such a complication obtained in this model.

Using formulas (29), (30), we can calculate the value of the day and night asymptotic temperatures for three planets: Mercury, Mars, and Pluto, calculating the parameter α (13) for them from the known experimental values: the planet's revolution period τ, the total radiation temperature due to solar radiation T_s, heat capacity c_v, density of matter on the surface of the planet ρ, and the depth of daily heating h [7]:

$$\alpha = \frac{\sigma \tau T_s^3}{2 c_v \rho h} \tag{33}$$

The following values were selected as average: heat capacity 1000 Joule/kg K, soil density 5000 kg /m^3, depth of daily heating of the planet's surface 1 m.

The results are presented in **Table 1**. It can be seen from it that the theoretical dimensional values of $T_{d\infty}$, $T_{n\infty}$ are in good agreement with their experimentally obtained values [7].

The asymptotic temperature ($t \to \infty$) of Mercury on the day side $T_{d\infty}$ is equal to the total radiation temperature due to solar radiation T_s. This is due to the fact that the day of Mercury is long enough; it manages to come into a state of thermal equilibrium with radiation. And during the night, it manages to cool noticeably: the temperature difference on the day and night sides is 480 K.

The days of Mars are relatively short. Therefore, the temperature difference is not so significant. The shorter the revolution period τ, the smaller the difference in the dimensional temperatures $T_{d\infty}$ and $T_{n\infty}$. In this regard, the difference in day and night temperatures in Mars is small and is equal to 10 K.

The length of the day in Pluto is several times longer than that of Mars, but the total radiation temperature due to solar radiation is much less due to the great remoteness of Pluto from the Sun. Because of this, its temperature practically does not change during its day.

7. Discussion of Results

In the world around, everything is ordered and expedient. Why so? Indeed, there

are always random collisions in it of the particles forming it, knocking out individual atoms from the ordered structures. If we didn't take any action that would return randomly detached atoms to their place, then the world would, according to the second law, surely collapse, and turn into chaos or cosmic dust in a finite time.

However, this does not happen. This means that there are guarding and restoring forces in nature, always returning "fallen atoms" to their place, *i.e.* ensuring the stability of the existence of each object: an elementary particle, the nucleus of an atom, cell, planet, star, ... the universe as a whole.

What are these forces doing? What does the "existence" of an object mean? When it's possible? Can forces of random, thermodynamic origin do this? Most likely, they cannot, and for a very simple reason. Because these forces (thermal conductivity, viscosity, diffusion, etc.) are rough, many micro particles always fall into their coverage area. But "building" and "restoring" forces need the ability to act locally; they influence on each particle separately in order to either build an ordered structured object in conditions of random collisions, or to recreate an object that deforms under random collisions. Otherwise, an attempt to rearrange, say, one micro particle, would entail the collective movement of many other micro particles that fell into the zone of action of thermodynamic forces.

8. Conclusions

The main result that can follow from this simple thermodynamic model of non-isolated systems: the cause of a decrease in entropy as a measure of disorder in such systems, which makes it possible for the phenomenon of "self-organization" to occur (the appearance of such structures as, say, Benard cells, the Belousov-Zhabotinsky reaction [8] [9] [10], etc.) is not the random process itself of entropy removal by the outgoing heat flow in non-isolated systems. As shown in this paper, it actually leads to the increase in entropy and to achieve a thermodynamic maximum in temperature and entropy in equilibrium state, as it takes in isolated systems.

An "entropy pump", which was regarded in some studies on synergetics as acting due to purely random processes of exporting entropy from a non-isolated system, does not work if we take into account the increase in entropy of the ingoing heat flow when it is thermalized at the input during the scattering of photons on the crystal lattice of the body (planet) surface.

The reason for the decrease in entropy can be regular factors of non-random nature (for example, periodic interruption of the ingoing heat flow), which can ensure that the temperature and entropy achieved in the new stationary state the values below the thermodynamic maximum.

Conflicts of Interest

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

References

[1] Prigogine (1955) Introduction to Thermodynamics of Irreversible Processes. Charles C Tomas, Springfield, IL.

[2] Ebeling, V., Engel, A. and Feistel, R. (2001) Physics of the Evolution Processes. M., URSS, 35+68.

[3] Prigogine, I. and Stengers, I. (1986) Order out of Chaos. A New Dialogue with Nature. M.: Progress, 432 p.

[4] Osipov, A.I. and Uvarov, A.V. (2004) *Soros Educational Journal*, **8**, 70.

[5] Cooney, F.M. (1981) Statistical Physics and Thermodynamics. M., Nauka.

[6] Landau, L.D. and Lifshits, E.M. (2002) Statistical Physics. Part I. V.5, M .; Fizmatlit, 616 p.

[7] Vorontsov-Veliaminov, B.A. and Straut, E.N. (2004) Drofa. Astronomy.

[8] Rudenko, A.P. (1969) The Theory of Self-Development of Non-Isolated Catalytic Systems. M.: Publishing House of Moscow State University, 276 p.

[9] Gershuni, G.Z. and Zhukhovitsky, E.M. (1972) Convective Instability of an Incompressible Fluid. M.: Nauka, 5.

[10] Prigogine, I. and Defay, R. (1966) Chemical Thermodynamics. Nauka. Siberian Branch, Novosibirsk.

Journal of Modern Physics, 2020, 11, 355-364
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Linking Conditions for Models with Geometrical Basis

Rainer Burghardt

A-2061 Obritz 246, Obritz, Austria
Email: arg@aon.at

How to cite this paper: Burghardt, R. (2020) Linking Conditions for Models with Geometrical Basis. *Journal of Modern Physics*, **11**, 355-364.
https://doi.org/10.4236/jmp.2020.113022

Received: February 11, 2020
Accepted: March 1, 2020
Published: March 4, 2020

Abstract

Usually it is demanded that the metric and its 1st derivatives have to match at the boundary of two adjacent regions which are solutions to Einstein's field equation. We propose a new linking condition concerning gravitational models based on surfaces which could be embedded into a higher dimensional flat space. We probe this condition for the Schwarzschild interior and exterior solution.

Keywords

Linking Conditions, Short Discussions of Previous Papers, Schwarzschild Models, Flamm' Paraboloid

1. Introduction

The question of how spaces with different geometrical structures can be adapted to each other takes up a lot of space in gravity theory. Stellar objects are described by interior solutions of Einstein's field equations, their gravitational fields by exterior solutions. The two solutions have to be adjusted at the surface of the stellar object. The constituent quantities of the two geometries must merge smoothly into one another. Numerous authors have dealt with the problem of junction conditions in recent years.

O'Brien and Synge [1] examined boundary conditions and jump conditions on surfaces where quantities and their derivatives can be discontinuous. To be consistent, they required the metrics and their 1st and 2nd derivatives to match at the boundary of two regions. Cocke [2] considered a non-static infinite cylinder that he cut out of a Friedman universe. The cylinder was surrounded by a gravitational field. For the linking condition for both regions, he relied on the metric and its 1st derivatives, which he treated as the 1st and 2nd fundamental forms.

Israel [3] [4] basically relied on the 2nd fundamental forms of a 3-surface. He discussed the physical discontinuities and mismatching of coordinates. He was mainly concerned with expanding spherical shells and their linking condition to the surrounding empty space. Bonnor [5] and Faulkes found a class of interior solutions that match an exterior solution with a moving boundary. As a linking condition, they used the matching of the metric and its 1st derivatives. For the interior solution, they used the interior Schwarzschild solution in isotropic coordinates. Lanczos [6] [7] considered in connection with the de Sitter cosmos two-dimensional distributed singularities in which the metrics remain finite and constant but take a jump with respect to the normals. He interpreted this as a surface distribution of matter. In another paper, he delved into the problem in more detail and replied to a criticism from Sen. He dealt in detail with the question to what extent the 1st derivatives of the metric must coincide at the boundary of two regions. Abraham [8] examined the discontinuities using the Gauss and Codazzi equations and builds on the generalized expressions of the O'Brien-Synge relations. The problem of matching two regions is also significant in cosmology. Galaxies and clusters are thought to be embedded into an FRW universe with homogeneous mass distribution. At the boundaries of such vacuoles, linking conditions must be adhered. We cite the paper of Gilbert [9] as a representative of this topic. Leibowitz [10] investigates junction conditions in going over to admissible coordinates in the case of comoving coordinates. He claims that the Oppenheimer-Snyder solutions are correctly matched. Attempted modifications are shown to be incorrect. Lichnerowicz [11] investigated junction conditions which can match up to the 3rd derivative of the functions at the boundary. Sen [12] described the discontinuities on a surface that is covered with matter. The examination was carried out independent of coordinate systems. Taub [13] faced the existence of 3-dimensional hyper surfaces in spacetime across which there may be discontinuities in the stress-energy-momentum tensor and the metric and their derivatives. Kumar [14] examined spherical shells in an empty universe. The 1st derivatives of the metric are discontinuous at the boundaries. The stress-energy-momentum tensor is defined with δ-functions. Coburn [15] determined discontinuity relations for a charged incompressible fluid with conservation laws and the 1st law of thermodynamics, and using shock waves. Edelen [16] obtained a dynamical theory of discontinuity surfaces and the associated jump strengths of both physical and geometrical quantities. It forms the basis for a general analysis of galactic structures. Huber [17] considered adjacent regions with different structures. He deformed the metrics of these regions in such a way that the linking conditions are satisfied at the boundary surface of these two regions. McVittie [18] studied collapsing models in a more general way and tried linking conditions as well. Dautcourt [19] [20] dealt with the jumps of the stress-energy-momentum tensor and considered layers on surfaces moving with the velocity of light. Papapetrou [21] and Treder investigated discontinuities on hypersurfaces and the associated problem of shock waves.

Hayward [22] discussed regions with boundary surfaces at which the normal unit vector changes discontinuously. The validity of the second linking condition was surveyed by Nariai and Tomita [23] [24] [25] for the collapsing Oppenheimer-Snyder model [26]. Although the metric of the interior OS solution matches the exterior OS solution—the Schwarzschild solution in comoving coordinates—the 1st derivatives of the metrics do not match at the boundary. Nariai and Tomita found a new exterior solution for the OS model, which fulfills the second linking condition for the OS interior and is free of singularity. Mitra [27] found that the two Schwarzschild solutions do not comply with the 2nd linking condition. He proposed a new interior solution that meets both linking conditions at the boundary to the exterior solution.

Thus, the methods of the Nariai, Tomita, and Mitra to solve the linking problem were quite different. We want to go a third way and replace the condition that the 1st derivatives of the metrics have to match with another that is quite plausible and that connects the interior and exterior Schwarzschild solutions.

In Sec. 2, we present the Schwarzschild geometry in the light of Flamm's [28] original paper. We focus on the radii of curvature of the normal and inclined slices of the surfaces on which the Schwarzschild geometry is based. The metrics have the signature 4. The time-like arc element is defined by $dx^4 = i(c)dt$. The tag "g" indicates the value of a quantity at the boundary of the surfaces. In Sec. 3, we show that the 1st derivatives of the metrics of the Schwarzschild models do not match, and we replace them by the postulate that the surfaces representing the interior and exterior Schwarzschild solutions have to have common tangents at the boundary surface.

2. The Schwarzschild Geometry

The new linking condition which we have introduced has a limited area of application. It can only be applied to models that can be explained geometrically, *i.e.*, models which have an embedding. We require that

I) the metrics match at the boundary.

II) the tangents (cutting tangents) of the embedded surfaces coincide.

We inspect this procedure facing the interior and exterior Schwarzschild solutions. Both solutions can be embedded into a 5-dimensional flat space, whereby a 6th variable is necessary for the exterior solution. The space-like part of the interior solution is represented by a spherical cap, the exterior part by Flamm's paraboloid. The two regions have to be matched.

Using quasi-polar coordinates, the standard form of the exterior Schwarzschild metric is formed as follows:

$$\mathrm{d}s^2 = \frac{1}{1-\frac{2M}{r}}\mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\mathrm{d}\varphi^2 - \left(1-\frac{2M}{r}\right)\mathrm{d}t^2. \qquad (2.1)$$

Here r is the radial coordinate. The space-like part of the metric is the line element on Flamm's paraboloid. The parabolic intersection curve of this surface,

i.e., the Schwarzschild parabola is given by

$$R^2 = 8M\left(r - 2M\right). \tag{2.2}$$

R is the coordinate of the extra dimension in the 5-dimensional embedding space normal to r. Flamm has given a detailed geometrical explanation. His proposed geometrical properties are shown in Figure 1.

Differentiating (2.2) and substituting for R, we obtain the ascent of the Schwarzschild parabola

$$\frac{\mathrm{d}R}{\mathrm{d}r} = \frac{4M}{R} = \frac{\sqrt{\frac{2M}{r}}}{\sqrt{1-\frac{2M}{r}}} = \tan\eta, \quad \sin\eta = \sqrt{\frac{2M}{r}}, \quad \cos\eta = \sqrt{1-\frac{2M}{r}}, \tag{2.3}$$

with the angle η as the angle of ascent of the Schwarzschild parabola.

A straight line normal to Flamm's paraboloid is cutting the coordinate R at the point P. The distance from P to the parabola is $\Re$ and the inclination η is the same as the angle of ascent of the parabola. From Figure 1, one can derive

$$r = \Re\sin\eta. \tag{2.4}$$

The radius of curvature of the Schwarzschild parabola can be calculated using elementary methods and (2.4)

$$\rho = \sqrt{\frac{2r^3}{M}} = 2r\sqrt{\frac{r}{2M}} = \frac{2r}{\sin\eta} = 2\Re.$$

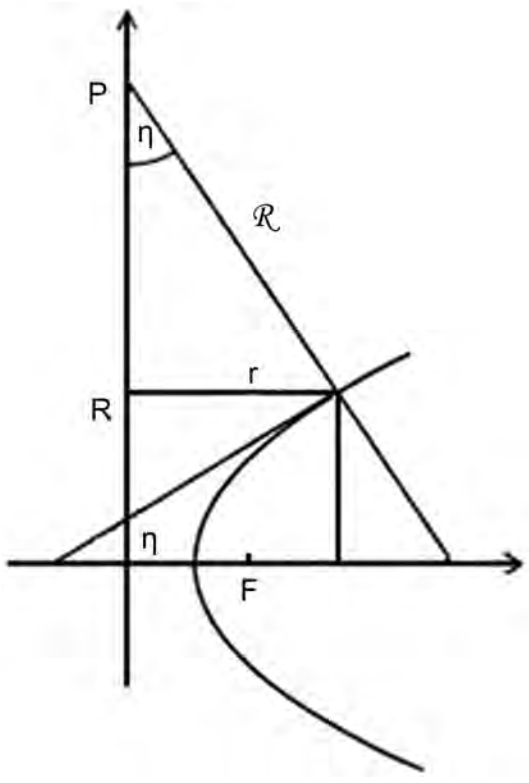


Figure 1. Flamm's explanation of the parabolic properties.

Finally, we find the fundamental relations

$$\rho = 2\Re, \quad \rho + \Re = 3\Re . \tag{2.5}$$

$3\Re$ is the distance between the point *P*, the "pole" of the parabola and the base point of the curvature vector of the parabola, lying on the evolute of the Schwarzschild parabola, *i.e.*, on Neil's parabola. With the insight of the factors 2 and 3, we have made a significant contribution to understand the interior Schwarzschild solution, as we will see later.

With the help of (2.3) the Schwarzschild metric can be written as

$$\mathrm{d}s^2 = \frac{1}{\cos^2\eta}\mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\mathrm{d}\varphi^2 - \cos^2\eta\mathrm{d}t^2 . \tag{2.6}$$

Further we put for the proper time

$$i\mathrm{d}T = \cos\eta\, i\mathrm{d}t = \rho\cos\eta\,\mathrm{d}i\psi . \tag{2.7}$$

Here, $i\psi$ is an imaginary angle and $\rho\cos\eta$ are the radii of a family of (open) pseudo circles (hyperbolae of constant curvature) lying in the $\left[x^{0'} = R, x^{4'}\right]$ -planes of the flat embedding space. $x^{4'}$ is an imaginary coordinate. A simple calculation shows that

$$\frac{1}{\cos\eta}\mathrm{d}r = -\rho\mathrm{d}\eta .$$

Thus, we are able to re-write the Schwarzschild line element as

$$\mathrm{d}s^2 = \rho_i\rho_k\mathrm{d}\eta^i\mathrm{d}\eta^k, \ \ i = k , \tag{2.8}$$

exhibiting all the curvatures of the slices of the surface described by the Schwarzschild metric. The curvature radii of the slices and the associated angles are

$$\begin{aligned} &\rho_1 = \rho = \sqrt{\frac{2r^3}{M}}, \quad \rho_2 = r, \quad \rho_3 = r\sin\vartheta, \quad \rho_4 = \rho\cos\eta \\ &\eta^1 = \eta, \quad \eta^2 = \vartheta, \quad \eta^3 = \varphi, \quad \eta^4 = i\psi. \end{aligned} \tag{2.9}$$

Tangents can be calculated on the intersection curves with these curvatures.

The line element of the interior Schwarzschild solution was given by Flamm as

$$\mathrm{d}s^2 = \Re^2\mathrm{d}\eta^2 + \Re^2\sin^2\eta\mathrm{d}\vartheta^2 + \Re^2\sin^2\eta\sin^2\vartheta\mathrm{d}\varphi^2 - \frac{1}{4}\left[3\cos\eta_g - \cos\eta\right]^2\mathrm{d}t^2 . \tag{2.10}$$

The space-like part of the metric is the metric of a hypersphere with the radius $\Re = const.$ and the polar angle η. But only a part of that hypersphere is used.

A spherical cap with the aperture angle η_g is cut off from the hypersphere and placed on the Schwarzschild parabola from below. The intersection curves of the spherical cap and the Schwarzschild parabola at the boundary must have common tangents (cutting tangents). We will write about this later. Unfortunately, Flamm's form of the interior metric is rarely found in literature. In the time-like part of the metric, the trigonometric functions are replaced by expressions with the radial variable *r*, by substituting $\sin\eta = r/\Re$. This avoids to un-

derstand the geometry.

Differentiating $r = \Re \sin\eta$, one finds

$$\Re \mathrm{d}\eta = \frac{1}{\cos\eta}\mathrm{d}r\,.$$

Writing for the coordinate time using (2.5), one obtains

$$i\mathrm{d}t = \rho_g \mathrm{d}i\psi = 2\Re \mathrm{d}i\psi\,, \tag{2.11}$$

where ρ_g is the curvature radius of the Schwarzschild parabola at the boundary surface. Finally, we have the equation for the metric of the interior Schwarzschild solution

$$\mathrm{d}s^2 = \frac{1}{1-\dfrac{r^2}{\Re^2}}\mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\mathrm{d}\varphi^2 + \left[3\Re\cos\eta_g - \Re\cos\eta\right]^2 \mathrm{d}i\psi^2\,. \tag{2.12}$$

Here, we have met the magical factors 2 and 3 as explained in (2.5). The proper time of the interior Schwarzschild model is described by two concentric pseudo-circles with the radii $3\Re\cos\eta_g$ and $\Re\cos\eta$. This circle is lying in the $\left[x^{0'}, x^{4'}\right]$-plane of the flat 5-dimensional embedding space[1].

We realize that the Schwarzschild interior solution is soldered to the exterior solution, because it contains elements of the exterior solution, *i.e.*, ρ_g the curvature radius of the Schwarzschild parabola at the boundary.

3. The Linking Conditions

We turn to the discussion of the linking conditions. Evidently, the 1st linking condition for the Schwarzschild solutions is satisfied. Their metrics coincide at the boundary. Considering the interior solution using (2.12) and putting $\eta = \eta_g$, one obtains

$$\frac{1}{\cos\eta_g}\mathrm{d}r, \quad 2\Re\cos\eta_g \mathrm{d}i\psi = \rho_g\cos\eta_g \mathrm{d}i\psi\,, \tag{3.1}$$

i.e., the corresponding expressions of the exterior metric (2.7) at the boundary surface.

It was Mitra [27], who showed that the 2nd linking condition cannot be applied to the Schwarzschild solutions. The first derivatives of the metrical coefficients do not match. The radial part of the line elements (2.6) and (2.12) for both solutions is

$$\frac{1}{\cos^2\eta}\mathrm{d}r^2\,.$$

Differentiating the metrical factor, we get

$$g_{11|1} = \left(\frac{1}{\cos^2\eta}\right)_{|1} = \frac{2\sin\eta}{\cos^3\eta}\eta_{|1}\,.$$

Here, all the indices are coordinate indices. Now one has to calculate $\eta_{|1}$ for both models. From

[1]More details one can find in our monographs [29] [30].

$$\frac{1}{\cos\eta}\mathrm{d}r = \Re\mathrm{d}\eta, \quad \frac{1}{\cos\eta}\mathrm{d}r = -\rho\mathrm{d}\eta$$

one has

$$\eta_{|1} = \frac{1}{\Re\cos\eta}, \quad \eta_{|1} = -\frac{1}{\rho\cos\eta}.$$

Thus, recalling $\rho = 2\Re$, one obtains the relations

$$g_{11|1} = \frac{2\sin\eta}{\Re\cos^4\eta}, \quad g_{11|1} = -\frac{2\sin\eta}{\rho\cos^4\eta} = -\frac{\sin\eta}{\Re\cos^4\eta}. \tag{3.2}$$

Evidently, these relations are also valid at the boundary surface. However, they differ by the factor −2, whereas the factor 2 is typically for the Schwarzschild geometry. Thus, the commonly accepted 2nd linking condition of O'Brien and Synge is not satisfied for the Schwarzschild geometry and has lost its legitimacy.

For the derivatives of the g_{44} one obtains for the interior solution

$$\begin{aligned} g_{44|1} &= \frac{1}{4}\left(3\cos\eta_g - \cos\eta\right)^2_{|1} \\ &= \frac{1}{2}\left(3\cos\eta_g - \cos\eta\right)\sin\eta\eta_{|1} \\ &= \frac{1}{2}\left(3\cos\eta_g - \cos\eta\right)\sin\eta\frac{1}{\Re\cos\eta} \end{aligned}$$

and at the boundary

$$g_{44|1} = \frac{1}{\Re}\sin\eta_g.$$

For the exterior solution one has

$$g_{44|1} = \left(\cos^2\eta\right)_{|1} = -2\cos\eta\sin\eta\eta_{|1} = -2\cos\eta\sin\eta\frac{1}{\rho\cos\eta} = -\frac{1}{\Re}\sin\eta, \tag{3.3}$$

and at the boundary

$$g_{44|1} = -\frac{1}{\Re}\sin\eta_g.$$

The difference in the signs can be explained with the fact that the curvature vectors $\mathfrak{R}$ and $\boldsymbol{\rho}$ have opposite directions. Robson [31] also recognizes that the 1st derivatives of the metric of the Schwarzschild models do not match at the boundary surface and he tries to force the match using a coordinate transformation. However, he drops this condition and agrees with the other authors to demand the coincidence of the 2nd fundamental forms.

In earlier papers we repeatedly mentioned that interior and exterior solutions should have common tangents at the boundary surface. We believed that this was a commonly accepted criterion for matching solutions. However, a careful study of the literature has shown that this requirement is not in use. In contrast, general validity is ascribed to the O'Brien-Synge method. Now we make up the proof that our requirement II provides functional results for the Schwarzschild geometry.

To calculate the tangents to the Flamm's paraboloid, it is sufficient to face the equation of the Schwarzschild parabola (2.2). We already have calculated the ascent of the Schwarzschild parabola with (2.3), *i.e.*,

$$\frac{\mathrm{d}R}{\mathrm{d}r} = \tan\eta\,, \tag{3.4}$$

which is equally valid at the boundary surface.

For the interior solution we have to calculate the ascent of the circle

$$r^2 + R^2 = \Re^2, \quad R = \pm\sqrt{\Re^2 - r^2}, \quad \frac{\mathrm{d}R}{\mathrm{d}r} = \pm\frac{r}{R}.$$

with $r = \Re\sin\eta, R = \Re\cos\eta$ one finally obtains

$$\frac{\mathrm{d}R}{\mathrm{d}r} = \tan\eta\,.$$

Here, the sign is chosen to be "+" because the spherical cap is adapted to the Schwarzschild parabola from below. We recognize that the interior and the exterior surfaces have a common tangent (cutting tangent) at the boundary surface. But this is evident right from the beginning, because the curvature vectors $\mathfrak{R}$ and $\boldsymbol{\rho}$ are lying in the same straight line at the boundary surface and are normal to the cap of the sphere and the Schwarzschild parabola and thus normal to the tangents of the two surfaces. Accordingly, the tangents have to coincide. We could have done without the calculation.

Lastly, we investigate the time-like parts of the models. Taking a glance at the interior metric (2.12), we find that the flow of time is characterized by two concentric pseudo circles with the radii $3\Re\cos\eta_g$ and $\Re\cos\eta$ founding a pseudo-ring sector. It is parameterized in the 5-dimensional flat space by

$$\begin{aligned} x^{0'} &= 3\Re\cos\eta_g\cos i\psi - \Re\cos\eta\cos i\psi \\ x^{4'} &= 3\Re\cos\eta_g\sin i\psi - \Re\cos\eta\sin i\psi \end{aligned}.$$

Since both circles have the same ascents, it is sufficient to calculate the ascent of one circle, *i.e.*, for a specific slice $\eta = const.$

$$\begin{aligned} &x^{0'^2} + x^{4'^2} = \Re^2\cos^2\eta \\ &\frac{\mathrm{d}x^{0'}}{\mathrm{d}x^{4'}} = -\frac{x^{4'}}{x^{0'}} = -\tan i\psi, \quad \frac{\mathrm{d}x^{0'}}{\mathrm{d}t'} = -i\tan i\psi = th\psi. \end{aligned} \tag{3.5}$$

For the exterior solution, one has

$$\begin{aligned} x^{0'} &= \rho\cos\eta\cos i\psi \\ x^{4'} &= \rho\cos\eta\sin i\psi \end{aligned}$$

and the equation of the pseudo circle for a specific slice $\rho\cos\eta = const.$ is

$$\begin{aligned} &x^{0'^2} + x^{4'^2} = \rho^2\cos^2\eta \\ &\frac{\mathrm{d}x^{0'}}{\mathrm{d}x^{4'}} = -\frac{x^{4'}}{x^{0'}} = -\tan i\psi, \quad \frac{\mathrm{d}x^{0'}}{\mathrm{d}t'} = -i\tan i\psi = th\psi \end{aligned}. \tag{3.6}$$

The result is also valid at the boundary surface. Thus, one gets common tangents.

Once more we note that the last calculation is superfluous since the expressions for the radii of the two interior pseudo circles reduce for $\eta = \eta_g$ to

$$2\Re \cos\eta_g = \rho_g \cos\eta_g .$$

The pseudo circles of the interior and exterior solutions coincide at the boundary surface and have the same tangents.

4. Conclusions

We showed that the interior Schwarzschild solution and the exterior Schwarzschild solution have common tangents at the boundary surface. We made this clear by calculating the ascents of the tangents of the two Schwarzschild solutions. Thus, the postulation that the tangents of surfaces representing gravitational models coincide at the boundary surface can serve as a linking condition and can replace the O'Brien-Synge condition, which does not apply to the Schwarzschild models.

Furthermore, it is quite likely that our method will also be applicable to the Reissner-Nordström model and to all models of the Kerr family.

Conflicts of Interest

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

References

[1] O'Brien, S. and Synge, J.L. (1952) *Communications of the Dublin Institute for Advanced Studies*, No. 9, 1-20.

[2] Cocke, W.J. (1966) *Journal of Mathematical Physics*, **7**, 1171-1178. https://doi.org/10.1063/1.1705020

[3] Israel, W. (1966) *Nuovo Cimento*, **44**, 1-14. https://doi.org/10.1007/BF02710419

[4] Israel, W. (1958) *Proceedings of the Royal Society of London. Series A*, **A248**, 404-414. https://doi.org/10.1098/rspa.1958.0252

[5] Bonnor, W.B. and Faulkes, M.C. (1967) *Monthly Notices of the Royal Astronomical Society*, **137**, 239-251. https://doi.org/10.1093/mnras/137.3.239

[6] Lanczos, K. (1922) *Annalen der Physik*, **24**, 539-543.

[7] Lanczos, K. (1924) *Annalen der Physik*, **74**, 518-540. https://doi.org/10.1002/andp.19243791403

[8] Abraham, R. (1962) *Journal of Mathematics and Mechanics*, **11**, 553-592. https://www.jstor.org/stable/24900913 https://doi.org/10.1512/iumj.1962.11.11033

[9] Gilbert, C. (1956) *Monthly Notices of the Royal Astronomical Society*, **116**, 678-683. https://doi.org/10.1093/mnras/116.6.678

[10] Leibowitz, C. (1969) *Nuovo Cimento*, **60B**, 254-260. https://doi.org/10.1007/BF02710226

[11] Lichnerowicz, A. (1971) *Comptes Rendus Chimie A*, **273**, 538.

[12] Sen, N. (1924) *Annalen der Physik*, **73**, 365-396.

https://doi.org/10.1002/andp.19243780505

[13] Taub, A.H. (1957) *Journal of Mathematical Physics*, **10**, 370-388. https://doi.org/10.1215/ijm/1255380389

[14] Kumar, M.M. (1970) *Progress of Theoretical Physics*, **44**, 243-253. https://doi.org/10.1143/PTP.44.243

[15] Coburn, N. (1961) *Journal of Mathematics and Mechanics*, **10**, 361-391. https://www.jstor.org/stable/24900727 https://doi.org/10.1512/iumj.1961.10.10023

[16] Edelen, D.G.B. (1963) *Journal of Mathematical Analysis and Applications*, **2**, 247-263. https://apps.dtic.mil/dtic/tr/fulltext/u2/407382.pdf https://doi.org/10.1016/0022-247X(63)90050-7

[17] Huber, A. (2019) Junction Conditions and Local Spacetimes in General Relativity.

[18] McVittie, G.E. (1964) *The Astrophysical Journal*, **140**, 401-416. https://doi.org/10.1086/147937

[19] Dautcourt, G. (1964) *Mathematische Nachrichten*, **27**, 279-288. https://doi.org/10.1002/mana.19640270504

[20] Dautcourt, G. (1963) *Archive for Rational Mechanics and Analysis*, **13**, 55-58. https://doi.org/10.1007/BF01262683

[21] Papapetrou, A. and Treder, H.-J. (1961) *Mathematische Nachrichten*, **23**, 371-384. https://doi.org/10.1002/mana.1961.3210230605

[22] Hayward, G. (1993) *Physical Review D*, **47**, 3275-3280. https://doi.org/10.1103/PhysRevD.47.3275

[23] Nariai, H. and Tomita, K. (1966) *Progress of Theoretical Physics*, **35**, 777-785. https://doi.org/10.1143/PTP.35.777

[24] Nariai, H. and Tomita, K. (1965) *Progress of Theoretical Physics*, **34**, 155-172. https://doi.org/10.1143/PTP.34.155

[25] Nariai, H. (1965) *Theoretical Physics*, **34**, 173-186. https://doi.org/10.1143/PTP.34.173

[26] Oppenheimer, J.R. and Snyder, H. (1939) *Physical Review*, **56**, 455-459. https://doi.org/10.1103/PhysRev.56.455

[27] Mitra, A. (2010) No Uniform Density Star in General Relativity. https://arxiv.org/pdf/1012.4985.pdf https://doi.org/10.1007/s10509-010-0567-8

[28] Flamm, L. (1916) *Physikalische Zeitschrift*, **17**, 448-454.

[29] Burghardt, R. (2016) Spacetime Curvature. 1-597. http://arg.or.at/EMono.htm

[30] Burghardt, R. (2016) Raumkrümmung. 1-623. http://arg.or.at/Mono.htm

[31] Robson, E.H. (1972) *Annales Henri Poincaré*, **16**, 41-50. http://eudml.org/doc/75725

Journal of Modern Physics, 2020, 11, 365-377
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Reduction of Superconducting Wave Packets in Dispersion Dynamics

Antony J. Bourdillon

UHRL, San Jose, CA, USA
Email: bourdillona@sbcglobal.net

How to cite this paper: Bourdillon, A.J. (2020) Reduction of Superconducting Wave Packets in Dispersion Dynamics. *Journal of Modern Physics*, **11**, 365-377.
https://doi.org/10.4236/jmp.2020.113023

Received: February 13, 2020
Accepted: March 6, 2020
Published: March 9, 2020

Abstract

Two problems in solid state physics and superconductivity are addressed by applications of dispersion dynamics. The first is the Hall effect. The dynamics of charges that yield positive Hall coefficients in material having no mobile positive charges have always been problematic The effect requires both electric and magnetic response, but magnetic deflection is only possible in mobile charges. In high temperature superconductors, these charges must be electrons. Contrary to Newton's second law, their acceleration is reversed in crystal fields that dictate negative dispersion. This is evident in room temperature measurements, but a second problem arises in supercurrents at low temperatures. The charge dynamics in material having zero internal electric field because of zero resistivity; and zero magnetic field because of the Meissner-Ochsenfeld diamagnetism; while the supercurrents themselves have properties of zero net momentum; zero spin; and sometimes, zero charge; are so far from having been resolved that they may never have been addressed. Again, dispersion dynamics are developed to provide solutions given by reduction of the superconducting wave packet. The reduction is here physically analyzed, though it is usually treated as a quantized unobservable.

Keywords

Reduction, Wave Packet, Dispersion Dynamics, Special Relativity, Propagation, Transverse Plane, Functions of Relativistic Free Particles, Quantum Physics, Quantum Mechanics

1. Introduction

Dispersion dynamics grows immediately out of Special relativity by the joint introduction of Planck's law and the de Broglie hypothesis. After first order differentiation of the standard formulae $E^2 = p^2c^2 + m_o^2c^4$ (where E represents

energy, p momentum, c the speed of light, and m_o rest mass) and the wave equivalent $\omega^2 = k^2c^2 + m_o^2c^4/\hbar^2$ (where ω represents angular frequency, k wave vector, and $\hbar$ the reduced Planck constant) [1]:

$$\frac{\mathrm{d}\omega}{\mathrm{d}k}\cdot\frac{\omega}{k} = v_g \cdot v_p = c^2 \tag{1}$$

i.e. the product of phase velocity v_p (angular frequency/wave vector) with group velocity v_g (the first derivative $\mathrm{d}\omega/\mathrm{d}k$) is equal to the square of the speed of light. Alternative normalization gives $v'_g \cdot v'_p = 1$, where $v'_g = v_g/c$ and $v'_p = v_p/c$. The group velocity is the speed of the particle in special relativity; the phase velocity is faster than light, and is measured either as the inverse of the group velocity or as the ratio of energy to momentum[1]. The phase velocity is singular in the rest frame: $v_p \to \infty$ when $k \to 0$. Massless particles *in vacuo* have $v_p = v_g$. More generally, $v_p = \left(1 + m_o^2/k^2\right)^{1/2}$, in units $c = \hbar = 1$; and $v_g = k/\left(k^2 + m_o^2\right)^{1/2} \simeq k/\omega$ when $k \ll m_o$, and $m_o \simeq \omega$.

Previously we have explored implications for Equation (1) in many fields including wave packets [2]; antiparticles in force fields [3]; electromagnetic interactions in wave groups [4]; the stable wave packet in quantum mechanics [5]; and in uncertainty [6] that is further developed below; voids in the Hall effect [7] and excitons in high temperature superconductors; many-body-gravitational solutions for galactic rotational velocities [8] that obviate dark matter and dark energy; magnetic radius in intrinsic spin that is equal to the Compton wavelength [9], etc. These applications spring clean old cobwebs and provide many cases of improved clarity for the way forward. Now we turn our attention to the reduction of the wave packet which is implicit in all calculations that involve transitions between time-independent wave functions, and which is particularly implicit in superconductivity.

Mathematicians choose their axioms; physicists falsify them. Typically, we investigate a new problem with simplicity, expecting complexity to later resolve oversimplification. As an early example, Newton believed light is corpuscular, moving in straight lines, where the complexity of wave-particle duality, came to dominate the subsequent legacy of 19th and 20th Century physics. The corresponding theoretical postulate of point particles[2], however convenient for limited purposes, poses anomalies for the real world of physics. For example, the electron may be regarded variously as a propagating wave in the electron microscope; or as an atomic wave orbital; or as an intrinsic magnetic current loop [9]; or having electrostatic radius. Each of these has dimensions over widely different orders of magnitude [10]. In particle physics, further dimensions occur in meson and hadron particles, and in their constituent quarks and interacting bosons.

Even at the relatively large scale of the atom, simplification is employed in

[1]The group and phase velocities supposedly combine to produce Dirac's "jitter" [Dirac, P.A.M., *The Principles of Quantum Mechanics* (1958) 4th Edition, Oxford] without purportedly offending the requirement for infinite energy in his electron particle, that he claims to have velocity *c*.

[2]Dirac supposed point particles apparently because the wave packet was, to him, unstable [ibid.].

time independent, wave functional solutions, where time is re-introduced through perturbations. The wave function then becomes a probability amplitude for locating an extreme point particle. This feature becomes more mysterious when it is reduced to reality during quantum measurement. Mystery and physics are antithetical. In particular, suppose a particle, say a photon, is emitted by an atom near point A so that it may contemporaneously arrive near B or C with appropriate probabilities on the same wave front. Two anomalous possibilities occur: on the one hand ask, "Can the particle be absorbed at both B and C by breaking energy conservation and doubling energy"; or on the other, "If absorption occurs at B, how does information arrive at C—without breaking the group velocity, v, constraints in special relativity ($v \leq c$, the speed of light)—in order to prevent a second absorption and to conserve energy?" A new solution for the reduction of the wave packet is available to Dispersion Dynamics [1] where the wave phase and particle group in the duality are treated specifically, having various properties that were previously ignored in standard quantum theory.

In the following discussion we suppose an *objective reality* in our mathematical framework that is not limited by Bohr's usage of *phenomena* [11]: "In contrast to his view that the notion of phenomenon *irrevocably* includes the specifics of the conditions of experimental observation, Einstein held that one should seek for a deeper-lying theoretical framework which permits the description of phenomena independently of these conditions"[3].

Before proceeding to a treatment of reduction, it will be necessary to clarify the notion of uncertainty in all its variations through multiple dimensions. The noteworthy example in superconductivity requires prior treatment of the Hall effect. In particular, high temperature superconductors, besides being ionic, have critical temperatures T_c that are typically an order of magnitude greater than classical, metallic, low temperature superconductors. Above respective critical temperatures $T > T_c$, the former group displays positive Hall coefficients, $R_H > 0$; the latter negative. The Hall voltage is generated in two stages: by a transient acceleration of electrons, across an applied magnetic field; in which magnetic deflection in the Lorentz force builds a charge potential in the steady state. The consequent field is normal to the two applied fields and produces a corresponding Hall voltage. When charge carriers are positive ions the Hall coefficient, $R_H > 0$; typically, negative charges yield $R_H < 0$. In high temperature superconductors, the charge carriers produce $R_H > 0$, though the only positive charges are statio-

[3]Pais [4] wrote that "It was [Einstein's] almost solitary conviction that quantum mechanics is logically consistent, but that it is an incomplete manifestation of an underlying theory in which an underlying objective reality is possible." Actually, in his EPR paper [Einstein, A., Podolsky, B. and Rosen, N., Phys.Rev. **47** 777-780 (19350] Einstein held that Bohr's theory is incomplete. The latter had claimed that *all* that can be known about an electron is in its wave function, particularly with regard to momentum and position. This seems an unlikely theory since Gödel's mathematical theorems on completeness and consistency in axiomatic systems [Gödel, K., *Monatshefte für Mathematik und Physik* 38 173-198 doi: 10.1007/BF01700692]. Popper was another realist in common with Einstein [Popper, K.R., *Quantum theory and the schism in physics*, Ed.Bartley W.W. III, (1982) Hutchinson]. This footnote does not contribute to the debate so much as outline background for arguments given in the text.

nary ions. These charges are subject to negative dispersion in ionic crystal fields [1] [7]. The true motion is described in dispersion dynamics.

A second and worse anomaly is the common description of superconducting currents. These are due to Cooper pairs of electrons that co-exist, in the superconducting state, with normal electrons [12]. In the superconductor not only is $\boldsymbol{E} = 0$ because of zero resistance; so also is magnetic field force $\boldsymbol{B} = 0$ because of the Meissner-Ochsenfeld effect (type II superconductors are slightly more complicated); while the net momentum carried by a Cooper pair $K = 0$; and likewise $\boldsymbol{S} = 0$. Moreover, in the high temperature superconductor, the excitonic net electric charge $\boldsymbol{Q} = 0$ [1]. Yet in spite of all of these absences, a supercurrent is observed. How can this be and what are the charge dynamics? Dispersion dynamics is used to show how the wave packet is reduced to produce the external current.

2. Uncertainty

In mathematical quantum mechanics the wave function is a probability amplitude for finding a point particle that is bounded by limits given by the Heisenberg uncertainty principle. By contrast, classical wave mechanics contains many techniques for predicting wave behaviors that depend on particular circumstances: the image plane; or near field; or far field; whether longitudinal or transverse; whether 1-dimensional; or 2-dimensional; or spherically symmetric; or cylindrical etc. Physical uncertainty describes properties of wave groups [2] where the alternative idea of a "limit" is less specific, and less useful to experimentalists, than "*expected uncertainty*" [5]. This is peculiar to each of the model systems to be described. We will consider first, the direction of propagation and then the transverse direction. in which the longitudinal wave packet is directly used. Waves are bounded in various ways. Waves that are not bounded have vanishing amplitude. Bounding in either space or time increases uncertainty in wave vector or angular frequency respectively, and *vice versa*.

2.1. Longitudinal

The stable wave packet ϕ is defined by mean angular frequency $\bar{\omega}$ and mean wave vector $\bar{k}$, which together stabilize the packet through conservation laws in energy and momentum (**Figure 1**):

$$\phi = A \cdot \exp\left(\frac{X^2}{2\sigma^2} + X\right), \text{ with } X = i\left(\bar{k}x - \bar{\omega}t\right) \tag{2}$$

where σ is the coherence, A is the wave amplitude, and x, t are coordinates in space and time. We can define the uncertainty in time by the full width at height $1/e$ ([3] p. 11):

$$\Delta t = 2\frac{\sigma}{\bar{\omega}} \tag{3}$$

and uncertainty of frequency by Fourier transform:

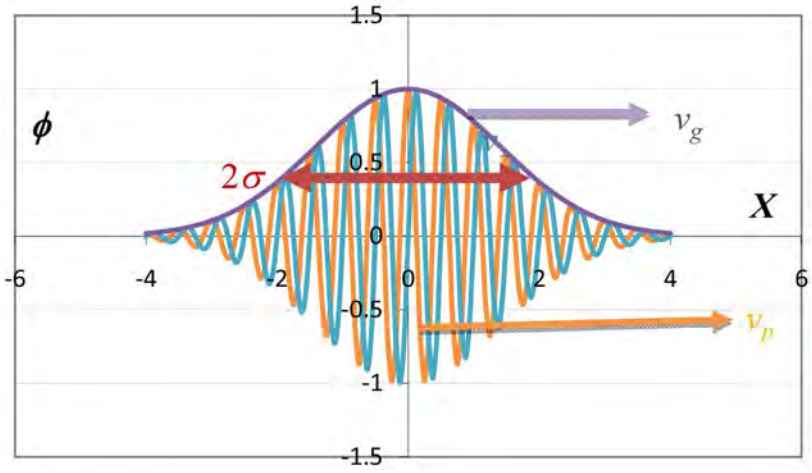


Figure 1. Stable wave packet (Equation (2)) containing envelope with group velocity v_g (purple arrow) and real (orange) and imaginary (blue) parts of the carrier wave having phase velocity v_p (orange arrow). The full width half maximum at $\phi = 1/e$, $2\sigma = \Delta t\bar{\omega}/2 = 4\bar{\omega}/\Delta\omega = \Delta x\bar{k}/2 = 4\bar{k}/\Delta k$ is shown by the red double arrow.

$$\Delta\omega = 4\frac{\bar{\omega}}{\sigma} \tag{4}$$

The product is a dual uncertainty in which σ cancels! By substituting for ω and using Planck's law:

$$\Delta E \cdot \Delta t = 8\hbar\,, \tag{5}$$

and by similar operations on de Broglie's hypothesis:

$$\Delta p_x \cdot \Delta x = 8\hbar\,. \tag{6}$$

Both values are more than an order of magnitude greater than Heisenberg's limit, ($\hbar/2$) (**Figure 2**). We shall see how they compare with uncertainty in the transverse plane. However, notice that limits are limitless, and valueless when they are too far away. In solid state theory the wave function is usually represented as indefinitely extended; but if extension is not a finite σ, the amplitude would be quantized zero.

2.2. Transverse 1-D in the Far Field

Next consider diffraction in the transverse plane that is caused by a horizontal slit, width d. The diffracted beam intensity tends to zero when the angle away from axis $\theta = \tan^{-1}(\lambda/d)$ where λ represents wavelength, and θ is sufficiently small. Then, using the de Broglie relationship, $p_x = h/\lambda$ and $\Delta p_y = 2p_x \sin(\theta) = 2h/d$ ([13] p.321), where h is Planck's constant.

$$\Delta p_y \cdot \Delta y = 2h\,. \tag{7}$$

This value is again greater than Heisenberg's limit, here by the factor 8π.

2.3. Transverse 1-D in the Near Field

It is well known that in Fresnel diffraction a narrow slit pinches a transmitted

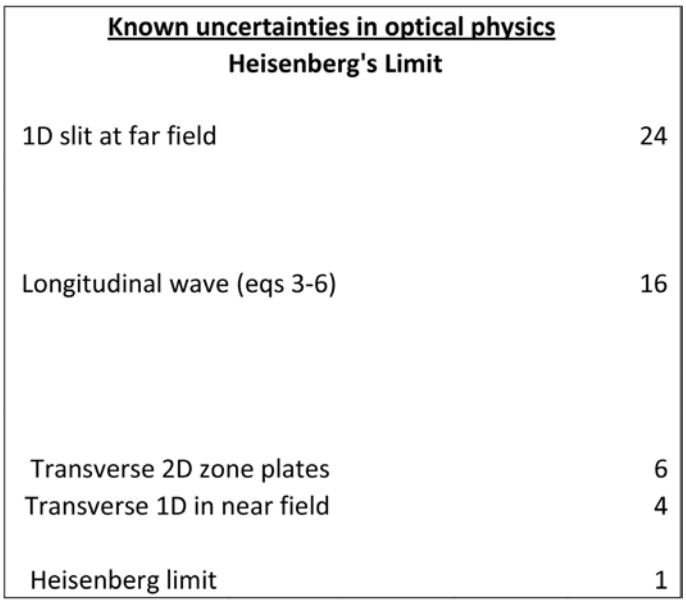


Figure 2. Uncertainties known in optical physics compared as multiples of Heisenberg's Limit.

beam in the near field, which is contrary to general effects elsewhere [13]. A zone in the wave field that has been used in X-ray lithography is called the Critical Condition [14] [15] [16] [17]. At this transverse plane, at a distance $\sim d^2/(3\lambda)$ downstream from the slit, the beam width is $\sim d/3$, From this it follows that $\Delta p_y = h/\lambda \cdot d/3 \cdot 3\lambda/d^2 = h/d$, and:

$$\Delta p_y \cdot \Delta y \sim h/3 . \tag{8}$$

The dual uncertainty is 6 times smaller than in Equation (7) and over 4 times larger than Heisenberg's limit. However, at gaps away from Critical, the contracted beam dimension varies rapidly and is illustrated elsewhere [13].

2.4. Transverse 2-D in Zone Plates

When a circular opaque screen of diameter Δs, is placed transverse to a parallel beam and viewed from downstream, an illuminated spot is seen on the axis. The intensity of the spot is greatly amplified when the circle is divided into optical half zones, of dimensions such that only zones of either odd or even sequence transmit. This zone plate acts like a lens [13], and aerial images show how beam profile varies with Gap d [18]. A typical dual uncertainty when the number of half zones is $N_F = (\Delta s)^2/(4d\lambda)$, is given by $\Delta p_y \cdot \Delta y = \Delta p_z \cdot \Delta z = h/2$, *i.e.* 4 times smaller than Equation (7) and 2π times greater than Heisenberg's limit. The varieties of uncertainty outlined by specific instances is needed in realistic applications of reduction in wave packets. To understand reduction, a limit of uncertainty is insufficient; only the expected value will do and this will vary with circumstances as already shown by the examples here in Section 2. The latter two cases are particularly extreme, yet are both significantly greater than Heisenberg's limit which is an inaccurate short cut. Moving forward, we must under-

stand the electronic "hole" before applying uncertainty to superconductivity.

3. The ELectronic "Hole" in Dispersion Dynamics

The problem we address is this: how is a hole deflected in a magnetic field? This is supposed to occur in the Hall effect measured in a p-type semiconductor, *P* doped *Si* for example. The charge carrier has a positive Hall coefficient like a positive ion such as a proton or Li^+ ion in whatever matrix. In the two step process, the charge is first accelerated in an electric field and then deflected in a crossed magnetic field. The force $\boldsymbol{F}$ on the deflection depends on the charge and is proportional to its velocity, v_g. Including the first stage electric acceleration. Instantaneously:

$$\boldsymbol{F} = q\boldsymbol{E} + q\boldsymbol{v}_g \times \boldsymbol{B} \tag{9}$$

Where q represents the carrier charge, now presumed electronic, negative e; $\boldsymbol{E}$ represents the applied Coulomb field; and $\boldsymbol{B}$ the crossed magnetic intensity. Since v_g is parallel to $\boldsymbol{E}$, the magnetic part $\boldsymbol{F}_m$ is normal to both $\boldsymbol{E}$ and $\boldsymbol{B}$ (Figure 3). The velocity is produced by the electric field, and can be averaged. For the more common case of the n-type semiconductor:

$$\boldsymbol{F} = -e\boldsymbol{E} - e\boldsymbol{v}_g \times \boldsymbol{B} \tag{10}$$

So that the acceleration a can be written approximately:

$$a \simeq -\frac{e\boldsymbol{E}}{m} - \frac{e\boldsymbol{E}t'}{m} \times \boldsymbol{B}\,. \tag{11}$$

where t' is a kind of mean time between collisions or the inverse of the number of collisions per unit time (Table 1)[4].

Since the only positive charges in *p*-type *Si*, for example, are on immobile nuclei, it is obvious that the carriers cannot be ions; they must be electrons. To understand how the charges react like positive ions to produce positive Hall coefficients, we have to return to Einstein's special relativity and derive the second derivative following Equation (1):

$$\frac{\mathrm{d}^2\omega}{\mathrm{d}k^2} = \frac{\mathrm{d}v_g}{\mathrm{d}k} = \frac{\mathrm{d}v_g}{\mathrm{d}p} = \left(\frac{1}{m'} - \frac{k^2}{m'^3}\right) = \frac{1}{m_{eff}} = \frac{a}{F} \tag{12}$$

where m' is the relativistic mass $m_o \Big/ \left(1 - v_g^2/c^2\right)^{1/2}$; effective mass m_{eff} is as defined in the brackets; a is acceleration in Newton's second law of motion corresponding to applied force F, such as the electric force (Equation (9)). in electricity and magnetism. Notice that a negative second derivative, or curvature (in *p-Si*, *Al* etc.), causes negative effective mass and negative acceleration under the *Coulomb force*. By contrast, the magnetic force is normal to the velocity and applied magnetic intensity so that its potential is structured differently from those in electrostatics and gravity. The result is that in magnetism, $\delta\omega = 0$ and the effective mass responding to the magnetic part of the electromagnetic Lorentz force is conventional (see Figure 3):

[4]This description has the same Hall result as previously [1] [7], but equations 10 and 11 had been misapplied to the magnetic field instead of the electric field.

Table 1. Commonly, the "hole" is supposed to operate as a positive charge. But the Hall effect in specimens where all positive charges are immobile can only have negatively charged carriers. These are not influenced by magnetic fields. However, following dispersion dynamics, (especially columns 2 and 4) the mobile electrons, as illustrated in **Figure 3**, are influenced by applied Coulomb forces and Lorentz magnetism. These are tabulated in the double event, namely electric acceleration and magnetic deflection.

	Charge q	$D^{\pm}$ curvature	m_{eff}	F_z^{el}	$a_z^{el} = F_z/m$	s	$\langle v_{gz} \rangle \simeq at'/2$	F_x^m	$a_x^m = F_x/m$	$q \cdot a_x^m \simeq E_x^{Hall}$
+ve ion	e	>0	m_i	eE	eE/m	>0	$eEt'/2m$	$eB(eEt'/2m)$	$eB(eEt'/m^2)$	>0
Electron	$-e$	>0	m	$-eE$	$-eE/m$	<0	$-eEt'/2m$	$eB(eEt'/2m)$	$eB(eEt'/m^2)$	<0
Holey electron	$-e$	<0	$-m$	$-eE$	eE/m	>0	$eEt'/2m$	$-eB(eEt'/2m)$	$-eB(eEt'/m^2)$	>0

Glossary: m_i ionic mass; a_z^{el} electric acceleration; t' time between collisions: expected velocity is half this time x acceleration; $\langle v_{gz} \rangle$ mean z-component velocity.

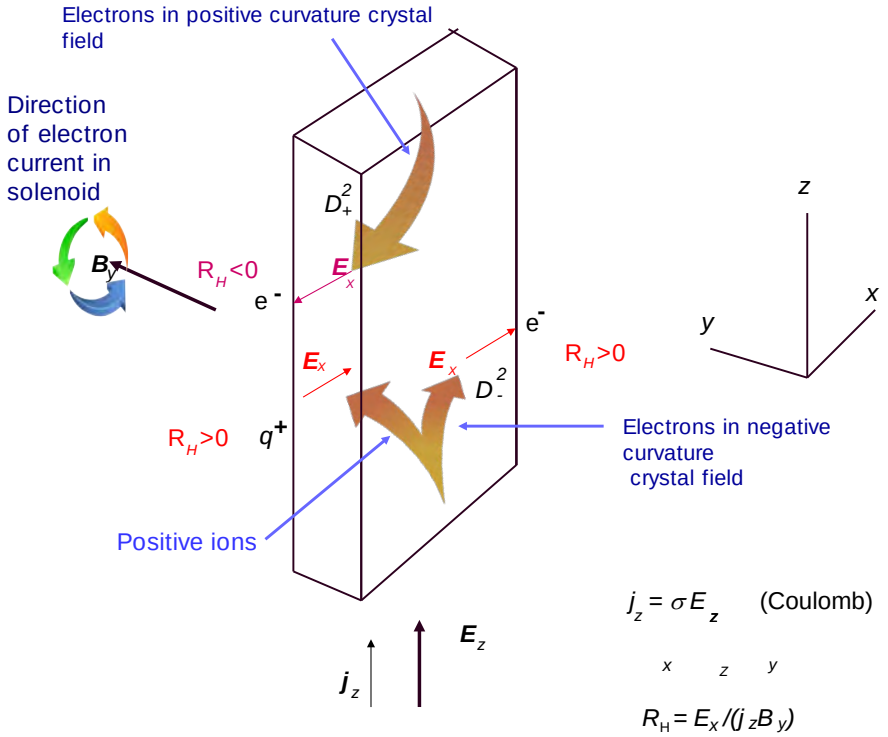


Figure 3. In the Hall effect, an applied electric field E_z is crossed with an applied magnetic field B_y. The resulting transient current is deflected and the consequential steady state charge causes a transverse electric field E_x. Equations (9)-(12) show how R_H is positive for mobile ions (lower left brown arrow) for p-type "holey" electrons (lower right); but more generally negative for n-type electrons (upper arrow). In dispersion dynamics, the current density j_z of p-type electrons reverses owing to negatively dispersed crystal fields.

$$\frac{\mathrm{d}v_g}{\mathrm{d}k} = \frac{1}{m} = \frac{a_x}{F_x^m}. \tag{13}$$

These formulae describe the motion of a single electron in an electron gas; they should be multiplied by carrier density to find measured values.

In summary:

1) The Hall effect occurs in two moments: a linear Coulomb acceleration; and

a magnetically deflected charge. The Hall voltage results from the steady state deflection of charges.

2) As for free electrons, in typical metals (*e.g.Cu*) the normal dispersion band curvature $D^{+} = \mathrm{d}^2\omega/\mathrm{d}k^2 > 0$ at Fermi level intercepts. $R_H < 0$ when the carriers have charge *-e* (Equation (11)).

3) By contrast, because all positive charges in *p-Si* and *Al* are immobile and since $R_H > 0$, the charge carriers in the magnetic field must be electrons.

4) Dispersion dynamics predicts that in crystal fields with negative dispersion D^{-}, the effective mass and acceleration are negative (Equation (12)).

5) In consequence, the velocity of electrons due to the Coulomb force F^{el} is reversed in D^{-}, compared with D^{+}.

6) In further consequence, the backward velocity reverses the otherwise conventional magnetic force F^{m} and deflection (**Table 1**, **Figure 3**). Thus in D^{-} crystal field dispersion, the Hall coefficient $R_H > 0$.

This gives the sign of R_H for "holey" electrons. Dispersion dynamics gives the reasons how and why those electrons behave like positive ions in the Hall effect. The argument is significant in high temperature superconductivity, where, because all positive charges are nuclear and immobile, the charge carriers can only be electrons [1] [7]. Indeed, charge pairing and reduction are common themes across the broad scope of bosonic wave particles in physics.

4. Reduction of Wave Packets in Massive and Massless Particles

Physically, how does it happen that a supercurrent is measured outside a superconductor, when its internal measurables, *E*, *B*, pair momentum *K*, pair spin *S*, resistivity *r*, etc., are all nominally zero? Reduction is mysterious when introduced as a change between time-independent wave functions: by hypothesis, the change may be represented as instantaneous at some point in space. In reality it is not so: in modern physics, resonant wave packets are the typical instruments for the force of change, and the packet is extended in both space and time. The space is divided between a direction of propagation and the corresponding transverse plane. Each of these differentiates as in Equation (1), where the *group* propagates with properties that are distinct from its internal transverse *phase.* In massive particles the group travels slower than light; the phase faster. The phase velocity v_P, in a particle with non-zero rest mass, is singular in the rest frame. Transmission of information is therefore theoretically instantaneous and Newtonian, *i.e.* within the transverse spatial coherence of the wave packet.

In massless photons travelling in vacuo, by contrast, the group and phase velocities share common speed *c*. In fact of course, the transmission is spread over time that is determined by the coherence 2σ, of the wave packet. This coherence determines likewise the spatial transmission, consistently with the methods of Huygen's wavelet construction and with the methods of Fresnel, Fraunhoferetc. [13].

Transmission of energy or momentum by *massive* particles, like the electron, is most easily considered in the rest frame: energy and momentum are conserved when a wave front encounters two targets simultaneously. The transverse phase velocity is singular and time is Newtonian, so that energy supplied to one target cannot be given to the other owing to released wave amplitude.

However, creation and annihilation of a *massless* particle, like the photon, is not so simple because c is the constant velocity for both group and phase velocities. By ignoring diffraction, the transverse spatial coherence then becomes c multiplied by the temporal coherence. This photon is described in Maxwell's theory by two real, sinusoidal, wave functions on mutually normal planes, one lagging in phase by $\pi/2$. Reduction of the packet thus occurs within its spatio-temporal coherence. From Equation (2), the *longitudinal* coherence length is found to be $2c\sigma$; the coherence time $2\sigma/c$. Through σ, the coherence values depend on initial conditions for the transmitting wave packet.

What can be understood about the *transverse* spatial coherence? Can a wave front become wider than the coherence? Does the coherence limit beam spread? Compare the temporal coherence from section 2.1 ($2\sigma/c$) with the transverse spatial coherence from Section 2.2 ($\tan^{-1}(\lambda/d)$). It is clear that when, typically $\lambda \ll d$, the ratio $c\Delta t/(\lambda/d) \sim 2\sigma d/\lambda$. Thus, when $\sigma > d/\lambda$, the transverse spatial coherence is determined by longitudinal temporal coherence; otherwise wave diffraction, d/λ, determines the transverse coherence. These features are in principle demonstrated by comparing diffraction of light generated by a laser source or point source. Supercurrents are massive, so we can now consider their reduction in their rest frame, *i.e.* in Newtonian time.

5. Super Currents

The solution for superconductivity was given by Bardeen, Cooper and Schrieffer [19] [20] [21]. Electrons in metallic superconductors occupy two states, normal and superconducting, *i.e.* at temperatures below critical $T < T_c$, and applied magnetic field $H < H_c$. The superconducting state consists of a pair of electrons bound together by a lattice distortion and energy gap ΔE. This Cooper pair has zero net momentum $K = 0$; zero net spin $S = 0$[5]; and electronic charge $2e$. The wave is superconductive, with zero resistivity $\rho = 0$ and therefore zero internal electric field $\boldsymbol{E} = 0$. Meanwhile zero internal magnetic field induction $\boldsymbol{B} = 0$ is demonstrated by the Meissner-Ochsenfeld effect. This is the basic state, though minor complexity arises in type II superconductors where electrons propagate quantized magnetic flux lines, and also at normal state boundaries as in the Josephson dc and ac tunneling effects. Since the normal and superconducting

[5]Even the intrinsic spin is problematic in dynamics. Spin is only measured in the presence of a magnetic field, so it is not obvious whether a measured magnetic moment is induced or "intrinsic" [1] [9]. Nor is it obvious that the electron is a point particle: it has magnetic moment with dimensions that include area? In particular, how can an "intrinsic spin" exist in the absence of a circulating current? Moreover, what physical transition is involved when a spin flip occurs? In dispersion dynamics, the spin is induced in non-resistive phase currents.

states are, by supposition, time independent, anomalies naturally arise in their dynamics. In high temperature superconductors, we understand the Cooper pair to be Wannier excitons with charge $Q = 0$ [1]. This is an added obstacle for conductivity. It suggests a real wave function that is Bosonic and commensurate with the lattice. How does transmission of current occur?

Consider for contrast, the more comprehensible oscillator strength in spectral transitions. This is a calculation of the relative transition rate between two stable electronic states, initial and final, excited by an operator O: $\left|\left\langle y_f \left| O \right| y_i \right\rangle\right|^2$, typically integrated over space, time, and density of states. In a dipole transition for example, $O = e\boldsymbol{r}$, where e represents electronic charge. Apply such a transition to measurement of resistivity of a superconductor by the four probe technique: at $T < T_c$ and $H < H_c$. When a voltage is applied between two outer terminals the supercurrent is measured at the inner sensor terminals. This current consists of normal electric charges that emerge from the positive terminal of a superconductor containing no electromagnetic fields. In dispersion dynamics this feature is not surprising: the terminals supply energy that breaks the Cooper pairs which release their normal charges to metallic terminals in Newtonian time, *i.e.* when considered in the rest frame. The dynamics will be governed by the transfer of thermal or electrical energy to the Cooper pairs. Each pair has zero net momentum, while any dynamic momentum that is transferred to the normal electrons at terminals, can be obtained by the pairs from the massive lattice as in X-ray diffraction and in the Mossbauer effect. Between the terminals, a supercurrent flows in absence from any internal electric or magnetic fields.

6. Summary and Conclusion

In dispersion dynamics, the product of the group velocity and phase velocity in a wave packet is equal to the square of the speed of light. In consequence, uncertainty is not an arbitrary limit, but a calculated expectation that has varied effects in reduction. Moreover, since immobile positive ions are not deflected by the magnetic Lorentz force, the positive Hall coefficients that are measured in certain metals and doped semiconductors must be generated by (negatively charged) "holey" electrons that exist in states having peculiar dispersive dynamics. Furthermore, in ionic high temperature superconductors, those electrons are supposedly contained in bosonic pairs of Wannier excitons. How then do supercurrents flow in the absence of internal electromagnetic fields and with zero momentum, zero spin, and zero net charge? In dispersion dynamics, the flow occurs by the reduction of time-independent waves that are consistent with packet decay, *i.e.* in Newtonian time between terminals. This reduction is a new solution for an unanswered problem. Following the analysis, the reduction is in principle verifiable by time-dependent measurements. Superconductivity is a physical paradigm for Boson statistics, for condensation, and for related phenomena that occur throughout the broad scope of physic. These further applications of dispersion dynamics will be objects for further study.

Conflicts of Interest

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

References

[1] Bourdillon, A.J. (2017) Dispersion Dynamics in the Hall Effect and Pair Bonding in HiT_c. Nova Science, Hauppauge.

[2] Bourdillon, A.J. (2012) *Journal of Modern Physics*, **3**, 290-296. https://doi.org/10.4236/jmp.2012.33041

[3] Bourdillon, A.J. (2013) *Journal of Modern Physics*, **4** 705-711. https://doi.org/10.4236/jmp.2013.46097

[4] Bourdillon, A.J. (2014) *Journal of Modern Physics*, **5**, 23-28. https://doi.org/10.4236/jmp.2014.51004

[5] Bourdillon, A.J. (2015) *Journal of Modern Physics*, **6**, 463-471. https://doi.org/10.4236/jmp.2015.64050

[6] Bourdillon, A.J. (2015) *Journal of Modern Physics*, **6**, 2011-2020. https://doi.org/10.4236/jmp.2015.614407

[7] Bourdillon, A.J. (2017) *Journal of Modern Physics*, **8**, 483-499. https://doi.org/10.4236/jmp.2017.84031

[8] Bourdillon, A.J. (2018) *Journal of Modern Physics*, **9**, 1304-1316. https://doi.org/10.4236/jmp.2018.96079

[9] Bourdillon, A.J. (2018) *Journal of Modern Physics*, *Special Issue on Magnetic Field and Magnetic Theory*, **9**, 2295-2307. https://doi.org/10.4236/jmp.2018.913145

[10] Aguilar-Benitez, M., Barnett, R.M., Caso, C., Conforto, G., Crawford, R.L., Cutkosky, R.E., Eichler, R.A., Eidelman, S., Groom, D.E., Hagiwara, K., Hayes, K.G., Hernandez, J.J., Hikasa, K., Höhler, G., Kawabata, S., Manley, D.M., Montanet, L., Morrison, R.J., Oliver, K.A., Porter, F.C., Roos, M., Schindler, R.H., Shrock, R.E., Stone, J., Törnqvist, N.A., Trippe, T.G., Wohl, C.G., Yost, G.P., Armstrong, B., Gieselmann, K. and Wagman, G.S. (1992) Particle Properties Data Booklet, June 1992. American Institute of Physics, from the Review of Particle Properties. Physical Review Data, D45.

[11] Pais, A. (1991) Niels Bohr's Times. Clarendon Press, Oxford, 433.

[12] Rose-Innes, A.C. and Rhoderick, E.H. (1978) Introduction to Superconductivity. 2nd Edition, Pergamon Press, Oxford.

[13] Jenkins, F.A. and White, E.W. (1937) Fundamentals of Optics. 4th Edition, McGraw-Hill, New York.

[14] Vladimirsky, Y., Bourdillon, A.J., Vladimirsky, O., Jiang, W. and Leonard, Q. (1999) *Journal of Physics D*: *Applied Physics*, **32**, L114-L118. https://doi.org/10.1088/0022-3727/32/22/102

[15] Bourdillon, A.J. and Vladimirsky, Y. (2005) X-Ray Lithography on the Sweet Spot.

[16] Bourdillon, A.J. and Vladimirsky, Y. (2002) Ultra High Resolution Lithographic Imaging and Printing by Exposure near the Critical Condition. US Patent 6383697.

[17] Bourdillon, A.J. (2005) Mask Shaping in Ultra High Resolution Lithography Using Temporal and Spatial Coherence. U.S. Patent 6953643.

[18] Bourdillon, A.J. and Boothroyd, C.B. (2005) *Journal of Physics D*: *Applied Physics*, **38**, 2947-2951. https://doi.org/10.1088/0022-3727/38/16/031

[19] Cooper, L.N. (1956) *Physical Review*, **104**, 1189-1190. https://doi.org/10.1103/PhysRev.104.1189

[20] Bardeen, J., Cooper, L.N. and Schrieffer, J.R. (1957) *Physical Review*, **106**, 162-164. https://doi.org/10.1103/PhysRev.106.162

[21] Bardeen, J., Cooper, L.N. and Schrieffer, J.R. (1957) *Physical Review*, **108**, 1175. https://doi.org/10.1103/PhysRev.108.1175

Journal of Modern Physics, 2020, 11, 378-394
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Generalization via Ultrahyperfunctions of a Gupta-Feynman Based Quantum Field Theory of Einstein's Gravity

A. Plastino[1,2,3], M. C. Rocca[1,2,4]

[1]Departamento de Física, Universidad Nacional de La Plata, Plata, Argentina
[2]Consejo Nacional de Investigaciones Científicas y Tecnológicas (IFLP-CCT-CONICET)-C. C. 727, La Plata, Argentina
[3]SThAR-EPFL, Lausanne, Switzerland
[4]Departamento de Matemática, Universidad Nacional de La Plata, Plata, Argentina
Email: angeloplastino@gmail.com

How to cite this paper: Plastino, A. and Rocca, M.C. (2020) Generalization via Ultrahyperfunctions of a Gupta-Feynman Based Quantum Field Theory of Einstein's Gravity. *Journal of Modern Physics*, **11**, 378-394.
https://doi.org/10.4236/jmp.2020.113024

Received: February 10, 2020
Accepted: March 6, 2020
Published: March 9, 2020

Abstract

Ultrahyperfunctions (UHF) are the generalization and extension to the complex plane of Schwartz' tempered distributions. This effort is an **application** to Einstein's gravity (EG) of the mathematical theory of convolution of Ultrahyperfunctions developed by Bollini *et al.* [1] [2] [3] [4]. A simplified version of these results was given in [5] and, based on them; a Quantum Field Theory (QFT) of EG [6] was obtained. Any kind of infinities is avoided by recourse to UHF. We will quantize EG by appealing to the most general quantization approach, the Schwinger-Feynman variational principle, which is more appropriate and rigorous that the popular functional integral method (FIM). FIM is not applicable here because our Lagrangian contains derivative couplings. We follow works by Suraj N. Gupta and Richard P. Feynman so as to undertake the construction of an EG-QFT. We explicitly use the Einstein Lagrangian as elaborated by Gupta [7], but choose a new constraint for the ensuing theory. In this way, we avoid the problem of lack of unitarity for the S matrix that afflicts the procedures of Gupta and Feynman. Simultaneously, we significantly simplify the handling of constraints, which eliminates the need to appeal to ghosts for guarantying unitarity of the theory. Our approach is obviously non-renormalizable. However, this inconvenience can be overcome by appealing to the mathematical theory developed by Bollini *et al.* [1] [2] [3] [4] [5]. Such developments were founded in the works of Alexander Grothendieck [8] and in the theory of Ultradistributions of Jose Sebastiao e Silva [9] (also known as Ultrahyperfunctions). Based on these works, an edifice has been constructed along two decades that are able to quantize non-renormalizable Field Theories (FT). Here we specialize this mathematical theory to discuss EG-QFT. Because we are using a Gupta-Feynman in-

spired EG Lagrangian, we are able to evade the intricacies of Yang-Mills theories.

Keywords

Quantum Field Theory, Einstein Gravity, Non-Renormalizable Theories, Unitarity

1. Introduction

Quantifying Einstein gravity (EG) remains an open question, a kind of supreme desideratum for quantum field theory (QFT). The failure of some attempts in this direction is due to the fact that 1) they appeal to Rigged Hilber Space (RHS) with undefined metric, 2) problems of non-unitarity, and also 3) non-renormalizablity issues.

Here we build up a unitary EG's QFT in the wake of related efforts by Suraj N. Gupta [7]. We deviate from his work by using a different EG-constraint, facing then a problem similar to that posed by Quantum Electrodynamics (QED). In order to quantize the concomitant non-renormalizable variational problem we appeal to mathematics developed by Bollini *et al.* [1] [2] [3] [4] [5], based upon the theory of Ultradistributions de J. Sebastiao e Silva (JSS) [9], also known as Ultrahyperfunctions (UHF). Any kind of infinities is avoided by recourse to UHF. The above cited mathematics was specifically devised to quantify non-renormalizable field theories. We consequently face a theory similar to QED, but endow with unitarity at all finite orders in power expansions in G (gravitation constant) of the EG Lagrangian. This was attempted without success first by Gupta and then by Feynman, in his Acta Physica Polonica work [10].

Mathematically, quantifying a non-renormalizable field theory is tantamount to suitably defining the product of two distributions (a product in a ring with zero-divisors in configuration space), an old problem in functional theory tackled successfully in [1] [2] [3] [4] [5].

Remarking that, in QFT, the problem of evaluating the product of distributions with coincident point singularities is related to the asymptotic behavior of loop integrals of propagators.

In references [1] [2] [3] [4], it was demonstrated that it is possible to define a general convolution between the ultradistributions of JSS [9] (Ultrahyperfunctions). This convolution yields another Ultrahyperfunction. Therefore, we have a product in a ring with zero divisors. Such a ring is the space of distributions of exponential type, or ultradistributions of exponential type obtained applying the anti-Fourier transform to the space of tempered ultradistributions or ultradistributions of exponential type.

We must clarify at this point that the ultrahyperfunctions, our present protagonists, are the generalization and extension to the complex plane of the Schwartz tempered distributions and the distributions of exponential type. That

is the tempered distributions and those of exponential type are a subset of the ultrahyprefunctions.

In our work we do not use counter-terms to get rid of infinities, because our convolutions are always finite. *We do not want counter-terms*, *since a non-renormalizable theory involves an infinite number of them*.

At the same time, we conserve all extant solutions to the problem of running coupling constants and the renormalization group. The convolution, once obtained, converts configuration space into a ring with zero-divisors. In it, one has now defined a product between the ring-elements. Thus, any unitary-causal-Lorentz invariant theory quantified in such a manner becomes predictive. The distinction between renormalizable on non-renormalizable QFT's becomes unnecessary now.

With our convolution, that uses Laurent's expansions (LE) in the parameter employed to define the LE, all finite constants of the convolutions become completely determined, eliminating arbitrary choices of finite constants. This is tantamount to eliminating all finite renormalizations of the theory. The independent term in the Laurent expansion yield the convolution value. This translates to configuration space the product-operation in a ring with divisors of zero.

We have already obtained an EG-OFT in [6] by recourse to the mathematics elaborated in [5]. What we do here is an extended EG-QFT, using the more general mathematical approach of [3].

The manuscript is organized as follows:

- Section 2 presents preliminary materials.
- In Section 3 we introduce the convolution of Ultrahyperfunctions, our present protagonists.
- Section 4 is devoted to the QFT Lagrangian for EG.
- In Section 5 we quantize the ensuing theory.
- In Section 6 the graviton's self-energy is evaluated up to second order.
- In Section 7 we introduce axions into our picture and deal with the axions gravitons interaction.
- In Section 8 we calculate the graviton's self-energy in the presence of axions.
- In Section 9 we evaluate, up to second order, the axion's self-energy.
- Finally, in Section 10, some conclusions are drawn.

2. Preliminary Materials

We do not deal in this effort with the popular functional integral method (FIM). Instead, we appeal here to the most general quantification approach, Schwinger-Feynman variational principle [11], which is able to deal even with high order supersymmetric theories, as exemplified by [12] [13]. Such theories cannot be quantized with the usual Dirac-brackets technique.

We introduce the action for a set of fields defined by

$$\mathcal{S}\left[\sigma(x),\sigma_0,\phi_A(x)\right]=\int_{\sigma_0}^{\sigma(x)}\mathcal{L}\left[\phi_A(\xi),\partial_\mu\phi_A(\xi),\xi\right]\mathrm{d}\xi, \tag{2.1}$$

where $\sigma(x)$ if a space-like surface passing through the point *x*. σ_0 is that surface at the remote past, at which all field variations vanish. The Schwinger-Feynman variational principle dictates that:

"Any Hermitian infinitesimal variation $\delta\boldsymbol{\mathcal{S}}$ of the action induces a canonical transformation of the vector space in which the quantum system is defined, and the generator of this transformation is this same operator $\delta\boldsymbol{\mathcal{S}}$".

Accordingly, the following equality holds:

$$\delta\phi_A = i\left[\delta\boldsymbol{\mathcal{S}}, \phi_A\right]. \tag{2.2}$$

Thus, for a Poincare transformation we have

$$\delta\boldsymbol{\mathcal{S}} = a^\mu \boldsymbol{\mathcal{P}}_\mu + \frac{1}{2} a^{\mu\nu} \boldsymbol{\mathcal{M}}_{\mu\nu}, \tag{2.3}$$

where the field variation is given by

$$\delta\phi_a = a^\mu \hat{P}_\mu \phi_A + \frac{1}{2} a^{\mu\nu} \hat{M}_{\mu\nu} \phi_A. \tag{2.4}$$

From (2.2) one gathers that

$$\partial_\mu \phi_A = i\left[\boldsymbol{\mathcal{P}}_\mu, \phi_A\right]. \tag{2.5}$$

Specifically,

$$\partial_0 \phi_A = i\left[\boldsymbol{\mathcal{P}}_0, \phi_A\right]. \tag{2.6}$$

This last result will be employed in quantizing EG.

3. The Convolution of Two Lorentz Invariant Tempered Ultradistributions

In [3] we have obtained a conceptually simple but rather lengthy expression for the convolution of two Lorentz invariant tempered ultradistributions:

$$\begin{aligned}
&H_\lambda(\rho,\Lambda) \\
&= \frac{1}{8\pi^2 \rho} \int_{\Gamma_1}\int_{\Gamma_2} F(\rho_1) G(\rho_2) \rho_1^\lambda \rho_2^\lambda \left\{ \Theta\left[\Im(\rho)\right] \left\{ \left[\ln(-\rho_1+\Lambda) - \ln(-\rho_1-\Lambda)\right] \right.\right. \\
&\times \left[\ln(-\rho_2+\Lambda) - \ln(-\rho_2-\Lambda)\right] \sqrt{4(\rho_1+\Lambda)(\rho_2+\Lambda) - (\rho-\rho_1-\rho_2-2\Lambda)^2} \\
&\times \ln\left[\frac{\sqrt{4(\rho_1+\Lambda)(\rho_2+\Lambda) - (\rho-\rho_1-\rho_2-2\Lambda)^2} - i(\rho-\rho_1-\rho_2-2\Lambda)}{2\sqrt{(\rho_1+\Lambda)(\rho_2+\Lambda)}} \right] \\
&+ \left[\ln(\rho_1+\Lambda) - \ln(\rho_1-\Lambda)\right]\left[\ln(\rho_2+\Lambda) - \ln(\rho_2-\Lambda)\right] \\
&\times \sqrt{4(\rho_1-\Lambda)(\rho_2-\Lambda) - (\rho-\rho_1-\rho_2+2\Lambda)^2} \\
&\times \ln\left[\frac{\sqrt{4(\rho_1-\Lambda)(\rho_2-\Lambda) - (\rho-\rho_1-\rho_2+2\Lambda)^2} - i(\rho-\rho_1-\rho_2+2\Lambda)}{2\sqrt{(\rho_1-\Lambda)(\rho_2-\Lambda)}} \right] \\
&+ \left[\ln(\rho_1+\Lambda) - \ln(\rho_1-\Lambda)\right]\left[\ln(-\rho_2+\Lambda) - \ln(-\rho_2-\Lambda)\right]
\end{aligned}$$

$$\times\left\{\frac{i\pi}{2}\left[\sqrt{4(\rho_1+\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)\right]\right.$$
$$+\sqrt{4(\rho_1+\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2)^2}$$
$$\left.\times\ln\left[\frac{\sqrt{4(\rho_1+\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)}{2i\sqrt{-(\rho_1+\Lambda)(\rho_2-\Lambda)}}\right]\right\}$$
$$+\left[\ln(-\rho_1+\Lambda)-\ln(-\rho_1-\Lambda)\right]\left[\ln(\rho_2+\Lambda)-\ln(\rho_2-\Lambda)\right]$$
$$\times\left\{\frac{i\pi}{2}\left[\sqrt{4(\rho_1-\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)\right]\right.$$
$$+\sqrt{4(\rho_1-\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2)^2}$$
$$\left.\left.\times\ln\left[\frac{\sqrt{4(\rho_1-\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)}{2i\sqrt{-(\rho_1-\Lambda)(\rho_2+\Lambda)}}\right]\right\}\right\}$$
$$-\Theta\left[-\Im(\rho)\right]\left\{\left[\ln(-\rho_1+\Lambda)-\ln(-\rho_1-\Lambda)\right]\left[\ln(-\rho_2+\Lambda)-\ln(-\rho_2-\Lambda)\right]\right.$$
$$\times\sqrt{4(\rho_1-\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2+2\Lambda)^2}$$
$$\times\ln\left[\frac{\sqrt{4(\rho_1-\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2+2\Lambda)^2}-i(\rho-\rho_1-\rho_2+2\Lambda)}{2\sqrt{(\rho_1-\Lambda)(\rho_2-\Lambda)}}\right]$$
$$+\left[\ln(\rho_1+\Lambda)-\ln(\rho_1-\Lambda)\right]\left[\ln(\rho_2+\Lambda)-\ln(\rho_2-\Lambda)\right]$$
$$\times\sqrt{4(\rho_1+\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2-2\Lambda)^2}$$
$$\times\ln\left[\frac{\sqrt{4(\rho_1+\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2-2\Lambda)^2}-i(\rho-\rho_1-\rho_2-2\Lambda)}{2\sqrt{(\rho_1+\Lambda)(\rho_2+\Lambda)}}\right]$$
$$+\left[\ln(\rho_1+\Lambda)-\ln(\rho_1-\Lambda)\right]\left[\ln(-\rho_2+\Lambda)-\ln(-\rho_2-\Lambda)\right]$$
$$\times\left\{\frac{i\pi}{2}\left[\sqrt{4(\rho_1-\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)\right]\right.$$
$$+\sqrt{4(\rho_1-\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2)^2}$$
$$\left.\times\ln\left[\frac{\sqrt{4(\rho_1-\Lambda)(\rho_2+\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)}{2i\sqrt{-(\rho_1-\Lambda)(\rho_2+\Lambda)}}\right]\right\}$$
$$+\left[\ln(-\rho_1+\Lambda)-\ln(-\rho_1-\Lambda)\right]\left[\ln(\rho_2+\Lambda)-\ln(\rho_2-\Lambda)\right]$$
$$\times\left\{\frac{i\pi}{2}\left[\sqrt{4(\rho_1+\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)\right]\right.$$
$$+\sqrt{4(\rho_1+\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2)^2}$$

$$\times \ln\left[\frac{\sqrt{4(\rho_1+\Lambda)(\rho_2-\Lambda)-(\rho-\rho_1-\rho_2)^2}-i(\rho-\rho_1-\rho_2)}{2i\sqrt{-(\rho_1+\Lambda)(\rho_2-\Lambda)}}\right]\Bigg\}\Bigg\}$$

$$-\frac{i}{2}\Big\{\big[\ln(-\rho_1+\Lambda)-\ln(-\rho_1-\Lambda)\big]\big[\ln(-\rho_2+\Lambda)-\ln(-\rho_2-\Lambda)\big]$$

$$\times(\rho_1-\rho_2)\left[\ln\left(i\sqrt{\frac{\rho_1+\Lambda}{\rho_2+\Lambda}}\right)+\ln\left(-i\sqrt{\frac{\rho_1-\Lambda}{\rho_2-\Lambda}}\right)\right]$$

$$+\big[\ln(\rho_1+\Lambda)-\ln(\rho_1-\Lambda)\big]\big[\ln(\rho_2+\Lambda)-\ln(\rho_2-\Lambda)\big]$$

$$\times(\rho_1-\rho_2)\left[\ln\left(-i\sqrt{\frac{\Lambda-\rho_1}{\Lambda-\rho_2}}\right)+\ln\left(i\sqrt{\frac{\Lambda+\rho_1}{\Lambda+\rho_2}}\right)\right]$$

$$+\big[\ln(\rho_1+\Lambda)-\ln(\rho_1-\Lambda)\big]\big[\ln(-\rho_2+\Lambda)-\ln(-\rho_2-\Lambda)\big]$$

$$\times\Bigg\{(\rho_1-\rho_2)\left[\ln\left(\sqrt{\frac{\Lambda+\rho_1}{\Lambda-\rho_2}}\right)+\ln\left(\sqrt{\frac{\Lambda-\rho_1}{\Lambda+\rho_2}}\right)\right]$$

$$+\frac{\rho_1-\rho_2}{2}\big[\ln(-\rho_1-\rho_2+\Lambda)-\ln(-\rho_1-\rho_2-\Lambda)$$

$$-\ln(\rho_1+\rho_2+\Lambda)+\ln(\rho_1+\rho_2-\Lambda)\big]+\rho_2\big[\ln(-\rho_1-\rho_2+\Lambda)$$

$$-\ln(-\rho_1-\rho_2-\Lambda)\big]+\rho_1\big[\ln(\rho_1+\rho_2+\Lambda)-\ln(\rho_1+\rho_2-\Lambda)\big]\Bigg\}$$

$$\big[\ln(-\rho_1+\Lambda)-\ln(-\rho_1-\Lambda)\big]\big[\ln(\rho_2+\Lambda)-\ln(\rho_2-\Lambda)\big]$$

$$\times\Bigg\{(\rho_1-\rho_2)\left[\ln\left(\sqrt{\frac{\Lambda-\rho_1}{\Lambda+\rho_2}}\right)+\ln\left(\sqrt{\frac{\Lambda+\rho_1}{\Lambda-\rho_2}}\right)\right]$$

$$+\frac{\rho_1-\rho_2}{2}\big[\ln(\rho_1+\rho_2+\Lambda)-\ln(\rho_1+\rho_2-\Lambda)$$

$$-\ln(-\rho_1-\rho_2+\Lambda)+\ln(-\rho_1-\rho_2-\Lambda)\big]+\rho_1\big[\ln(-\rho_1-\rho_2+\Lambda)\qquad(3.1)$$

$$-\ln(-\rho_1-\rho_2-\Lambda)\big]+\rho_2\big[\ln(\rho_1+\rho_2+\Lambda)-\ln(\rho_1+\rho_2-\Lambda)\big]\Bigg\}\Bigg\}\Bigg\}\mathrm{d}\rho_1\mathrm{d}\rho_2$$

This defines an ultradistribution in the variables ρ and Λ for

$$\left|\Im(\rho)\right|>\Im(\Lambda)>\left|\Im(\rho_1)\right|+\left|\Im(\rho_2)\right|$$

Let $\mathfrak{B}$ be a vertical band contained in the complex λ-plane $\mathfrak{P}$. Integral (3.1) is an analytic function of λ defined in the domain $\mathfrak{B}$. Moreover, it is bounded by a power of $|\rho\Lambda|$. Then, $H_\lambda(\rho,\Lambda)$ can be analytically continued to other parts of $\mathfrak{P}$. Thus, we define

$$H(\rho)=H^{(0)}(\rho,i0^+)\qquad(3.2)$$

$$H_\lambda(\rho,i0^+)=\sum_{-m}^{\infty}H^{(n)}(\rho,i0^+)\lambda^n\qquad(3.3)$$

As in the other cases, we define now

$$\{F*G\}(\rho)=H(\rho) \tag{3.4}$$

as the convolution of two Lorentz invariant tempered ultradistributions.

The Feynman propagators corresponding to a massless particle F and a massive particle G are, respectively, the following ultrahyperfunctions:

$$F(\rho)=-\Theta\left[-\Im(\rho)\right]\rho^{-1}$$

$$G(\rho)=-\Theta\left[-\Im(\rho)\right]\left(\rho+m^2\right)^{-1} \tag{3.5}$$

where ρ is the complex variable, such that on the real axis one has $\rho=k_1^2+k_2^2+k_3^2-k_0^2$. For them, the following equalities are satisfied

$$\rho^{\lambda}F(\rho)=-\Theta\left[-\Im(\rho)\right]\rho^{\lambda-1}$$

$$\rho^{\lambda}G(\rho)=-\Theta\left[-\Im(\rho)\right]\left(\rho+m^2\right)^{\lambda-1} \tag{3.6}$$

where we have used: $\left(\rho+m^2\right)^{\lambda}\simeq\rho^{\lambda}$, since we have chosen m to be very small. On the real axis, the previously defined propagators are given by:

$$f(\rho)=F(\rho+i0)-F(\rho-i0)=(\rho-i0)^{-1}$$

$$g(\rho)=G(\rho+i0)-G(\rho-i0)=\left(\rho+m^2-i0\right)^{-1} \tag{3.7}$$

These are the usual expressions for Feynman propagators.

Consider first the convolution of two massless propagators. We use (3.6), since here the corresponding ultrahyperfunctions do not have singularities in the complex plane. We obtain from (3.1) a simplified expression for the convolution:

$$h_{\lambda}(\rho)=\frac{\pi}{2\rho}\iint_{-\infty}^{\infty}(\rho_1-i0)^{\lambda-1}(\rho_2-i0)^{\lambda-1}\left[(\rho-\rho_1-\rho_2)^2-4\rho_1\rho_2\right]_{+}^{\frac{1}{2}}\mathrm{d}\rho_1\mathrm{d}\rho_2 \tag{3.8}$$

This expression is nothing other than the usual convolution:

$$h_{\lambda}(\rho)=(\rho-i0)^{\lambda-1}*(\rho-i0)^{\lambda-1} \tag{3.9}$$

In the same way, we obtain for massive propagators:

$$h_{\lambda}(\rho)=\left(\rho+m^2-i0\right)^{\lambda-1}*\left(\rho-m^2-i0\right)^{\lambda-1} \tag{3.10}$$

These last two expressions are the ones we will use later to evaluate the graviton's self-energy.

4. The Lagrangian of Einstein's QFT

Our EG Lagrangian reads [7]

$$\mathcal{L}_G=\frac{1}{\kappa^2}\boldsymbol{R}\sqrt{|g|}-\frac{1}{2}\eta_{\mu\nu}\partial_{\alpha}h^{\mu\alpha}\partial_{\beta}h^{\nu\beta}, \tag{4.1}$$

where $\eta^{\mu\nu}=diag(1,1,1,-1)$, $h^{\mu\nu}=\sqrt{|g|}g^{\mu\nu}$ The second term in (4.1) fixes the gauge. We effect now the linear approximation

$$h^{\mu\nu}=\eta^{\mu\nu}+\kappa\phi^{\mu\nu}, \tag{4.2}$$

where κ^2 is the gravitation's constant and $\phi^{\mu\nu}$ the graviton field. We write

$$\mathcal{L}_G = \mathcal{L}_L + \mathcal{L}_I, \tag{4.3}$$

where

$$\mathcal{L}_L = -\frac{1}{4}\Big[\partial_\lambda \phi_{\mu\nu}\partial^\lambda \phi^{\mu\nu} - 2\partial_\alpha \phi_{\mu\beta}\partial^\beta \phi^{\mu\alpha} + 2\partial^\alpha \phi_{\mu\alpha}\partial_\beta \phi^{\mu\beta}\Big], \tag{4.4}$$

and, up to 2nd order, one has [7]:

$$\mathcal{L}_I = -\frac{1}{2}\kappa\phi^{\mu\nu}\left[\frac{1}{2}\partial_\mu \phi^{\lambda\rho}\partial_\nu \phi_{\lambda\rho} + \partial_\lambda \phi_{\mu\beta}\partial^\beta \phi_\nu^\lambda - \partial_\lambda \phi_{\mu\rho}\partial^\lambda \phi_\nu^\rho\right], \tag{4.5}$$

having made use of the constraint

$$\phi_\mu^\mu = 0. \tag{4.6}$$

This constraint is required in order to satisfy gauge invariance [14] For the graviton we have then

$$\Box\phi_{\mu\nu} = 0, \tag{4.7}$$

whose solution is

$$\phi_{\mu\nu} = \frac{1}{(2\pi)^{\frac{3}{2}}}\int\left[\frac{a_{\mu\nu}(\boldsymbol{k})}{\sqrt{2k_0}}\mathrm{e}^{ik_\mu x^\mu} + \frac{a^+_{\mu\nu}(\boldsymbol{k})}{\sqrt{2k_0}}\mathrm{e}^{-ik_\mu x^\mu}\right]\mathrm{d}^3k, \tag{4.8}$$

with $k_0 = |\boldsymbol{k}|$.

5. Quantization of the Theory

We need some definitions. The energy-momentum tensor reads

$$T_\rho^\lambda = \frac{\partial\mathcal{L}}{\partial\partial^\rho\phi^{\mu\nu}}\partial^\lambda\phi^{\mu\nu} - \delta_\rho^\lambda\mathcal{L}, \tag{5.1}$$

and the time-component of the four-momentum is

$$\mathcal{P}_0 = \int T_0^0\mathrm{d}^3x. \tag{5.2}$$

Using (4.4) we have

$$\begin{aligned} T_0^0 = \frac{1}{4}\Big[&\partial_0\phi_{\mu\nu}\partial^0\phi^{\mu\nu} + \partial_j\phi_{\mu\nu}\partial^j\phi^{\mu\nu} - 2\partial_\alpha\phi_{\mu0}\partial^0\phi^{\mu\alpha} - 2\partial_\alpha\phi_{\mu j}\partial^j\phi^{\mu\alpha} \\ &+ 2\partial_\alpha\phi^{\mu\alpha}\partial_0\phi_\mu^0 + 2\partial_\alpha\phi^{\mu\alpha}\partial_j\phi_\mu^j\Big]. \end{aligned} \tag{5.3}$$

Consequently,

$$\mathcal{P}_0 = \frac{1}{4}\int|\boldsymbol{k}|\Big[a_{\mu\nu}(\boldsymbol{k})a^{+\mu\nu}(\boldsymbol{k}) + a^{+\mu\nu}(\boldsymbol{k})a_{\mu\nu}(\boldsymbol{k})\Big]\mathrm{d}^3k. \tag{5.4}$$

Appeal to (2.6) leads to

$$\begin{aligned} \Big[\mathcal{P}_0, a_{\mu\nu}(\boldsymbol{k})\Big] &= -k_0 a_{\mu\nu}(\boldsymbol{k}) \\ \Big[\mathcal{P}_0, a^{+\mu\nu}(\boldsymbol{k})\Big] &= k_0 a^{+\mu\nu}(\boldsymbol{k}). \end{aligned} \tag{5.5}$$

From the last relation in (5.5) one gathers that

$$|\boldsymbol{k}|a^{+\rho\lambda}(\boldsymbol{k}') = \frac{1}{2}\int|\vec{k}|\Big[a_{\mu\nu}(\boldsymbol{k}), a^{+\rho\lambda}(\boldsymbol{k}')\Big]a^{+\mu\nu}(\boldsymbol{k})\mathrm{d}^3k. \tag{5.6}$$

The solution of this integral equation is

$$\left[a_{\mu\nu}(\boldsymbol{k}),a^{+\rho\lambda}(\boldsymbol{k}')\right]=\left[\delta_{\mu}^{\rho}\delta_{\nu}^{\lambda}+\delta_{\nu}^{\rho}\delta_{\mu}^{\lambda}\right]\delta(\boldsymbol{k}-\boldsymbol{k}'). \tag{5.7}$$

As customary, the physical state $|\psi\rangle$ of the theory is defined via the equation

$$\phi_{\mu}^{\mu}|\psi\rangle=0. \tag{5.8}$$

We use now the usual definition

$$\Delta_{\mu\nu}^{\rho\lambda}(x-y)=\langle 0|T\left[\phi_{\mu\nu}(x)\phi^{\rho\lambda}(y)\right]|0\rangle. \tag{5.9}$$

The graviton's propagator then turns out to be

$$\Delta_{\mu\nu}^{\rho\lambda}(x-y)=\frac{i}{(2\pi)^4}\left(\delta_{\mu}^{\rho}\delta_{\nu}^{\lambda}+\delta_{\nu}^{\rho}\delta_{\mu}^{\lambda}\right)\int\frac{\mathrm{e}^{ik_{\mu}\left(x^{\mu}-y^{\mu}\right)}}{k^2-i0}\mathrm{d}^4k. \tag{5.10}$$

As a consequence, we can write

$$\mathcal{P}_0=\frac{1}{4}\int|\boldsymbol{k}|\left[a_{\mu\nu}(\boldsymbol{k})a^{+\mu\nu}(\boldsymbol{k}')+a^{+\mu\nu}(\boldsymbol{k}')a_{\mu\nu}(\boldsymbol{k})\right]\delta(\boldsymbol{k}-\boldsymbol{k}')\mathrm{d}^3k\mathrm{d}^3k', \tag{5.11}$$

or

$$\mathcal{P}_0=\frac{1}{4}\int|\boldsymbol{k}|\left[2a^{+\mu\nu}(\boldsymbol{k}')a_{\mu\nu}(\boldsymbol{k})+\delta(\boldsymbol{k}-\boldsymbol{k}')\right]\delta(\boldsymbol{k}-\boldsymbol{k}')\mathrm{d}^3k\mathrm{d}^3k'. \tag{5.12}$$

Thus, we obtain

$$\mathcal{P}_0=\frac{1}{2}\int|\boldsymbol{k}|a^{+\mu\nu}(\boldsymbol{k})a_{\mu\nu}(\boldsymbol{k})\mathrm{d}^3k, \tag{5.13}$$

where we have used the fact that the product of two deltas with the same argument vanishes [1], *i.e.*, $\delta(\boldsymbol{k}-\boldsymbol{k}')\delta(\boldsymbol{k}-\boldsymbol{k}')=0$. This illustrates the fact that using Ultrahyperfunctions is here equivalent to adopting the normal order in the definition of the time-component of the four-momentum

$$\mathcal{P}_0=\frac{1}{4}\int|\boldsymbol{k}|:\left[a_{\mu\nu}(\boldsymbol{k})a^{+\mu\nu}(\boldsymbol{k})+a^{+\mu\nu}(\boldsymbol{k})a_{\mu\nu}(\boldsymbol{k})\right]:\mathrm{d}^3k. \tag{5.14}$$

Now, we must insist on the fact that the physical state should satisfy not only Equation (5.8) but also the relation (see [7])

$$\partial_{\mu}\phi^{\mu\nu}|\psi\rangle=0. \tag{5.15}$$

The ensuing theory is similar to the QED-one obtained via the quantization approach of Gupta-Bleuler. This implies that the theory is unitary for any finite perturbative order. In this theory only one type of graviton emerges, ϕ^{12}, while in Gupta's approach two kinds of graviton arise. Obviously, this happens for a non-interacting theory, as remarked by Gupta.

Undesired Effects of NOT Using Our Constraint

If we do NOT use the constraint (5.8), we have

$$\mathcal{P}_0=\frac{1}{2}\int|\boldsymbol{k}|\left[a^{+\mu\nu}(\boldsymbol{k})a_{\mu\nu}(\boldsymbol{k})-\frac{1}{2}a_{\mu}^{+\mu}(\boldsymbol{k})a_{\nu}^{\nu}(\boldsymbol{k})\right]\mathrm{d}^3k, \tag{5.16}$$

and, appealing to the Schwinger-Feynman variational principle we find

$$|\boldsymbol{k}|a_{\rho\lambda}^{+}(\boldsymbol{k}')=\frac{1}{2}\int|\boldsymbol{k}|\left\{a^{+\mu\nu}(\boldsymbol{k})\left[a_{\mu\nu}(\boldsymbol{k}),a_{\rho\lambda}^{+}(\boldsymbol{k}')\right]-\frac{1}{2}a_{\mu}^{+\mu}(\boldsymbol{k})\left[a_{\nu}^{\nu}(\boldsymbol{k}),a_{\rho\lambda}^{+}(\boldsymbol{k}')\right]\right\}\mathrm{d}^3k, \tag{5.17}$$

whose solution is

$$\left[a_{\mu\nu}(\boldsymbol{k}),a^{+}_{\rho\lambda}(\boldsymbol{k}')\right]=\left[\eta_{\mu\rho}\eta_{\nu\lambda}+\eta_{\nu\rho}\eta_{\mu\lambda}-\eta_{\mu\nu}\eta_{\rho\lambda}\right]\delta(\boldsymbol{k}-\boldsymbol{k}'). \tag{5.18}$$

The above is the customary graviton's quantification, that leads to a theory whose S matrix in not unitary [7] [10].

6. The Self Energy of the Graviton

To evaluate the graviton's self-energy (SF)c we start with the interaction Hamiltonian $\mathcal{H}_I$. Note that the Lagrangian contains derivative interaction terms.

$$\mathcal{H}_I=\frac{\partial\mathcal{L}_I}{\partial\partial^0\phi^{\mu\nu}}\partial^0\phi^{\mu\nu}-\mathcal{L}_I. \tag{6.1}$$

A typical term reads

$$\Sigma_{G\alpha_1\alpha_2\alpha_3\alpha_4}(k)=k_{\alpha_1}k_{\alpha_2}(\rho-i0)^{\lambda-1}*k_{\alpha_3}k_{\alpha_4}(\rho-i0)^{\lambda-1}. \tag{6.2}$$

where $\rho=k_1^2+k_2^2+k_3^2-k_0^2$

The Fourier transform of (6.2) is

$$\begin{aligned}
&\mathcal{F}\left[k_{\alpha_1}k_{\alpha_2}(\rho-i0)^{\lambda-1}*k_{\alpha_3}k_{\alpha_4}(\rho-i0)^{\lambda-1}\right]\\
&=-\frac{2^{4(\lambda+1)}\left[\Gamma(2+\lambda)^2\right]}{4\Gamma(1-\lambda)^2}\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4}(x+i0)^{-2\lambda-4}\\
&+\frac{2^{4(\lambda+1)}\Gamma(2+\lambda)\Gamma(3+\lambda)}{2\Gamma(1-\lambda)^2}\left(\eta_{\alpha_1\alpha_2}x_{\alpha_3}x_{\alpha_4}+\eta_{\alpha_3\alpha_4}x_{\alpha_1}x_{\alpha_2}\right)(x+i0)^{-2\lambda-5}\\
&-\frac{2^{4(\lambda+1)}\Gamma(3+\lambda)^2}{\Gamma(1-\lambda)}x_{\alpha_1}x_{\alpha_2}x_{\alpha_3}x_{\alpha_4}(x+i0)^{-\nu-2}
\end{aligned} \tag{6.3}$$

where $x=x_1^2+x_2^2+x_3^2-x_0^2$

Anti-transforming the above equation we have

$$\begin{aligned}
&k_{\alpha_1}k_{\alpha_2}(\rho-i0)^{\lambda-1}*k_{\alpha_3}k_{\alpha_4}(\rho-i0)^{\lambda-1}\\
&=\frac{i\pi^2}{4\Gamma(1-\lambda)^2}\left\{\Gamma(\lambda+2)\left[\frac{\Gamma(2+\lambda)}{\Gamma(2\lambda+4)}-2\frac{\Gamma(3+\lambda)}{\Gamma(2\lambda+5)}\right]\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4}\right.\\
&\left.+\frac{\Gamma(\lambda+3)^2}{\Gamma(2\lambda+6)}\left(\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4}+\eta_{\alpha_2\alpha_3}\eta_{\alpha_1\alpha_4}+\eta_{\alpha_2\alpha_4}\eta_{\alpha_1\alpha_3}\right)\right\}\Gamma(-2\lambda-2)(\rho-i0)^{2\lambda+2}\\
&+\frac{i\pi^2\Gamma(\lambda+3)}{2\Gamma(1-\lambda)^2}\left\{\frac{\Gamma(2+\lambda)}{\Gamma(2\lambda+5)}\Gamma(\nu+1)\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4}+\eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2}\right)\right.\\
&-\frac{\Gamma(\lambda+3)}{\Gamma(2\lambda+6)}\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4}+\eta_{\alpha_1\alpha_3}k_{\alpha_2}k_{\alpha_4}+\eta_{\alpha_1\alpha_4}k_{\alpha_2}k_{\alpha_3}+\eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2}\right.\\
&\left.\left.+\eta_{\alpha_2\alpha_3}k_{\alpha_1}k_{\alpha_4}+\eta_{\alpha_2\alpha_4}k_{\alpha_1}k_{\alpha_3}\right)\right\}\Gamma(-2\lambda-1)(\rho-i0)^{2\lambda+1}\\
&+\frac{i\pi^2\Gamma(\lambda+3)^2}{\Gamma(1-\lambda)^2\Gamma(2\lambda+6)}k_{\alpha_1}k_{\alpha_2}k_{\alpha_3}k_{\alpha_4}\Gamma(-2\lambda)(\rho-i0)^{2\lambda}
\end{aligned} \tag{6.4}$$

Self-Energy Evaluation for $\lambda = 0$

We appeal now to a λ-Laurent expansion and retain there the $\lambda = 0$ independent term [3]. Thus, we Laurent-expand (6.4) around $\lambda = 0$ and find

$$
\begin{aligned}
& k_{\alpha_1}k_{\alpha_2}\left(\rho - i0\right)^{\lambda-1} * k_{\alpha_3}k_{\alpha_4}\left(\rho - i0\right)^{\lambda-1} \\
& = -i\frac{\pi^2}{4\lambda}\left\{\frac{1}{5!}\left(\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4} + \eta_{\alpha_2\alpha_3}\eta_{\alpha_1\alpha_4} + \eta_{\alpha_2\alpha_4}\eta_{\alpha_1\alpha_3}\right)\rho^2\right. \\
& -\left[\frac{2}{4!}\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4} + \eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2}\right) - \frac{1}{6!}\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4}\right.\right. \\
& + \eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2} + \eta_{\alpha_1\alpha_3}k_{\alpha_2}k_{\alpha_4} + \eta_{\alpha_1\alpha_4}k_{\alpha_2}k_{\alpha_3} \\
& \left.\left. + \eta_{\alpha_2\alpha_3}k_{\alpha_1}k_{\alpha_4} + \eta_{\alpha_2\alpha_4}k_{\alpha_1}k_{\alpha_3}\right)\right]\rho + \frac{8}{5!}k_{\alpha_1}k_{\alpha_2}k_{\alpha_3}k_{\alpha_4}\left.\right\} \\
& -\frac{i\pi^2}{5!2}\left(\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4} + \eta_{\alpha_2\alpha_3}\eta_{\alpha_1\alpha_4} + \eta_{\alpha_2\alpha_4}\eta_{\alpha_1\alpha_3}\right)\left[\ln\left(\rho - i0\right) - \frac{137}{60}\right]\rho^2 \\
& + i\frac{\pi^2}{4!}\left\{\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4} + \eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2}\right)\left[\ln\left(\rho - i0\right) - \frac{11}{6}\right]\right. \\
& -\frac{1}{24}\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4} + \eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2} + \eta_{\alpha_1\alpha_3}k_{\alpha_2}k_{\alpha_4} + \eta_{\alpha_1\alpha_4}k_{\alpha_2}k_{\alpha_3}\right. \\
& \left.\left. + \eta_{\alpha_2\alpha_3}k_{\alpha_1}k_{\alpha_4} + \eta_{\alpha_2\alpha_4}k_{\alpha_1}k_{\alpha_3}\right)\left[\ln\left(\rho - i0\right) + \ln\pi + 2C - \frac{101}{30}\right]\right\}\rho \\
& \left. - i\frac{\pi^2}{30}k_{\alpha_1}k_{\alpha_2}k_{\alpha_3}k_{\alpha_4}\left[\ln\left(\rho - i0\right) - \frac{47}{60}\right] + \sum_{n=1}^{\infty}a_n\lambda^n\right\}.
\end{aligned}
\tag{6.5}
$$

The exact value of the convolution we are interested in, *i.e.*, the left hand side of (5.5), is given by the independent term in the above expansion, as it is well-known. If the reader is not familiar with this situation, see for instance [3]. We then reach

$$
\begin{aligned}
& \Sigma_{G\alpha_1\alpha_2\alpha_3\alpha_4}\left(k\right) = k_{\alpha_1}k_{\alpha_2}\left(\rho - i0\right)^{-1} * k_{\alpha_3}k_{\alpha_4}\left(\rho - i0\right)^{-1} \\
& = -\frac{i\pi^2}{5!2}\left(\eta_{\alpha_1\alpha_2}\eta_{\alpha_3\alpha_4} + \eta_{\alpha_2\alpha_3}\eta_{\alpha_1\alpha_4} + \eta_{\alpha_2\alpha_4}\eta_{\alpha_1\alpha_3}\right)\left[\ln\left(\rho - i0\right) - \frac{137}{60}\right]\rho^2 \\
& + i\frac{\pi^2}{4!}\left\{\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4} + \eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2}\right)\left[\ln\left(\rho - i0\right) - \frac{11}{6}\right]\right. \\
& -\frac{1}{24}\left(\eta_{\alpha_1\alpha_2}k_{\alpha_3}k_{\alpha_4} + \eta_{\alpha_3\alpha_4}k_{\alpha_1}k_{\alpha_2} + \eta_{\alpha_1\alpha_3}k_{\alpha_2}k_{\alpha_4} + \eta_{\alpha_1\alpha_4}k_{\alpha_2}k_{\alpha_3}\right. \\
& \left.\left. + \eta_{\alpha_2\alpha_3}k_{\alpha_1}k_{\alpha_4} + \eta_{\alpha_2\alpha_4}k_{\alpha_1}k_{\alpha_3}\right)\left[\ln\left(\rho - i0\right) + \ln\pi + 2C - \frac{101}{30}\right]\right\}\rho \\
& \left. - i\frac{\pi^2}{30}k_{\alpha_1}k_{\alpha_2}k_{\alpha_3}k_{\alpha_4}\left[\ln\left(\rho - i0\right) - \frac{47}{60}\right]\right\}.
\end{aligned}
\tag{6.6}
$$

We have to deal with 1296 diagrams of this kind.

7. Including Axions into the Picture

Axions are hypothetical elementary particles postulated by the Peccei-Quinn theory in 1977 to tackle the strong CP problem in quantum chromodynamics. If

they exist and have low enough mass (within a certain range), they could be of interest as possible components of cold dark matter [15]. We include now a massive scalar field (axions) interacting with the graviton. The Lagrangian becomes

$$\mathcal{L}_{GM} = \frac{1}{\kappa^2}\boldsymbol{R}\sqrt{|g|} - \frac{1}{2}\eta_{\mu\nu}\partial_\alpha h^{\mu\alpha}\partial_\beta h^{\nu\beta} - \frac{1}{2}\left[h^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + m^2\phi^2\right]. \tag{7.1}$$

We can now recast the Lagrangian in the fashion

$$\mathcal{L}_{GM} = \mathcal{L}_L + \mathcal{L}_I + \mathcal{L}_{LM} + \mathcal{L}_{IM}, \tag{7.2}$$

where

$$\mathcal{L}_{LM} = -\frac{1}{2}\left[\partial_\mu\phi\partial^\mu\phi + m^2\phi^2\right], \tag{7.3}$$

so that $\mathcal{L}_{IM}$ becomes the Lagrangian for the axion-graviton action

$$\mathcal{L}_{IM} = -\frac{1}{2}\kappa\phi^{\mu\nu}\partial_\mu\phi\partial_\nu\phi. \tag{7.4}$$

The new term in the interaction Hamiltonian is

$$\mathcal{H}_{IM} = \frac{\partial\mathcal{L}_{IM}}{\partial\partial^0\phi}\partial^0\phi - \mathcal{L}_{IM}. \tag{7.5}$$

8. The Complete Self Energy of the Graviton

The presence of axions generates a new contribution to the graviton's self energy

$$\Sigma_{GM\,\mu rvs}(k) = k_\mu k_r\left(\rho + m^2 - i0\right)^{-1} * k_v k_s\left(\rho + m^2 - i0\right)^{-1}. \tag{8.1}$$

So as to compute it we appeal to the usual integral together with the generalized Feynman-parameters. After a Wick rotation we obtain

$$\begin{aligned} &k_\mu k_r\left(\rho + m^2 - i0\right)^{\lambda-1} * k_v k_s\left(\rho + m^2 - i0\right)^{\lambda-1} \\ &= i\int_0^1 x^{-\lambda}(1-x)^{-\lambda}\int\frac{k_\mu k_r(p_v - k_v)(p_s - k_s)}{\left[(k-px)^2 + a\right]^{2-2\lambda}}\mathrm{d}^4k\mathrm{d}x, \end{aligned} \tag{8.2}$$

where

$$a = p^2x - p^2x^2 + m^2. \tag{8.3}$$

After the variables-change $u = k - px$ we find

$$\begin{aligned} &k_\mu k_r\left(\rho + m^2 - i0\right)^{\lambda-1} * k_v k_s\left(\rho + m^2 - i0\right)^{\lambda-1} \\ &= i\int_0^1 x^{-\lambda}(1-x)^{-\lambda}\int\frac{f(u,x,\mu,r,v,s)}{\left(u^2 + a\right)^{2-2\lambda}}\mathrm{d}^4u\mathrm{d}x \end{aligned} \tag{8.4}$$

where

$$\begin{aligned} &f(u,x,\mu,r,v,s) \\ &= u_\mu u_r p_v p_s(1-x)^2 + u_\mu u_r u_v u_s - u_\mu u_s p_r p_v x(1-x) \\ &\quad - u_\mu u_v p_r p_s x(1-x) - u_r u_s p_\mu p_v x(1-x) \\ &\quad - u_r u_v p_\mu p_s x(1-x) + p_\mu p_r p_v p_s x^2(1-x)^2 + u_v u_s p_\mu p_r x^2. \end{aligned} \tag{8.5}$$

After evaluation of the pertinent integrals we arrive at

$$
\begin{aligned}
&k_\mu k_r\left(\rho+m^2-i0\right)^{\lambda-1} * k_\nu k_s\left(\rho+m^2-i0\right)^{\lambda-1} \\
&=\frac{i\pi^{\frac{5}{2}}2^{2\lambda}m^{2+4\lambda}}{16}\frac{\Gamma(-1-2\lambda)}{\Gamma(1-\lambda)}\left(\eta_{\mu r}k_\nu k_s+\eta_{\nu s}k_\mu k_r\right) \\
&\times\left[\frac{F\left(1-2\lambda,-1-2\lambda,\frac{3}{2}-\lambda;-\frac{\rho}{4m^2}\right)}{\Gamma\left(\frac{3}{2}-\lambda\right)}+\frac{F\left(1-\lambda,-1-2\lambda,\frac{5}{2}-\lambda;-\frac{\rho}{4m^2}\right)}{2\Gamma\left(\frac{5}{2}-\lambda\right)}\right] \\
&+i\frac{i\pi^{\frac{5}{2}}2^{2\lambda-1}m^{4+4\lambda}}{4}\left(\eta_{\mu r}\eta_{\nu s}+\eta_{\mu\nu}\eta_{rs}+\eta_{\mu s}\eta_{\nu r}\right) \\
&\frac{\Gamma(-2-2\lambda)}{\Gamma(1-\lambda)\Gamma\left(\frac{3}{2}-\lambda\right)}F\left(-2-2\lambda,1-\lambda,\frac{3}{2}-\lambda;-\frac{\rho}{4m^2}\right) \\
&-i\frac{i\pi^{\frac{5}{2}}2^{2\lambda}m^{2+4\lambda}}{64}\left(\eta_{\mu s}k_r k_\nu+\eta_{\mu\nu}k_r k_s+\eta_{rs}k_\mu k_\nu+\eta_{r\nu}k_\mu k_s\right) \\
&\times\frac{\Gamma(2-\lambda)\Gamma(-1-2\lambda)}{\Gamma(1-\lambda)^2\Gamma\left(\frac{5}{2}-\lambda\right)}F\left(-1-2\lambda,2-\lambda,\frac{5}{2}-\lambda;-\frac{\rho}{4m^2}\right) \qquad (8.6) \\
&+i\frac{i\pi^{\frac{5}{2}}2^{2\lambda}m^{4\lambda}}{32}k_\mu k_r k_\nu k_s\frac{\Gamma(3-\lambda)\Gamma(-2\lambda)}{\Gamma(1-\lambda)^2\Gamma\left(\frac{5}{2}-\lambda\right)}F\left(-2\lambda,2-\lambda,\frac{5}{2}-\lambda;-\frac{\rho}{4m^2}\right).
\end{aligned}
$$

Self-Energy Evaluation for $\lambda=0$

We need again a Laurent's expansion and face

$$
\begin{aligned}
&k_\mu k_r\left(\rho+m^2-i0\right)^{\lambda-1} * k_\nu k_s\left(\rho+m^2-i0\right)^{\lambda-1} \\
&=i\frac{\pi^2}{4\lambda}\left\{m^2\left(\eta_{\mu r}k_\nu k_s+\eta_{\nu s}k_\mu k_r\right)\left[\frac{1}{3}+\frac{1}{5}\frac{\rho}{4m^2}\right]\right. \\
&-m^4\left(\eta_{\mu r}\eta_{\nu s}+\eta_{\mu\nu}\eta_{rs}+\eta_{\mu s}\eta_{r\nu}\right)\times\left[\frac{1}{4}+\frac{1}{3}\frac{\rho}{4m^2}+\frac{4}{15}\left(\frac{\rho}{4m^2}\right)^2\right] \\
&-\frac{m^2}{4m^2+k^2-i0}\left(\eta_{\mu s}k_r k_\nu+\eta_{\mu\nu}k_r k_s+\eta_{rs}k_\mu k_\nu+\eta_{r\nu}k_\mu k_s\right) \\
&\left.\times\frac{k^2-m^2}{12}+\frac{m^2}{4}+\frac{k^2-m^2}{30}\frac{\rho}{4m^2}-\frac{1}{6}k_\mu k_r k_\nu k_s\right\} \\
&+i\frac{m^2\pi^2}{2}\left(\eta_{\mu r}k_\nu k_s+\eta_{\nu s}k_\mu k_r\right) \\
&\times\left[\frac{1}{3}\left(\ln m^2+\frac{1}{12}\right)+\frac{1}{5}\frac{\rho}{4m^2}\left(\ln m^2+\frac{13}{15}\right)\right] \\
&+i\frac{m^2\pi^2}{30}\left(\eta_{\mu r}k_\nu k_s+\eta_{\nu s}k_\mu k_r\right)\frac{\rho}{4m^2} \\
&\times\left[F\left(1,1,\frac{7}{2};-\frac{\rho}{4m^2}\right)+\frac{1}{7}F\left(1,1,\frac{9}{2};-\frac{\rho}{4m^2}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&-i\frac{\pi^2 m^4}{4}\left(\eta_{\mu r}\eta_{vs}+\eta_{\mu v}\eta_{rs}+\eta_{\mu s}\eta_{vr}\right)\\
&\times\left\{\left[\frac{1}{2}-\frac{2}{3}\frac{\rho}{4m^2}-\frac{8}{15}\left(\frac{\rho}{4m^2}\right)^2\right]\right.\\
&\left.\times\left(\ln m^2+1\right)-\frac{1}{2}\left[\frac{3}{2}-\frac{1}{9}\left(\frac{\rho}{4m^2}\right)+\frac{52}{225}\left(\frac{\rho}{4m^2}\right)^2\right]\right\}\\
&-i\frac{2\pi^2 m^4}{105}\left(\eta_{\mu r}\eta_{vs}+\eta_{\mu v}\eta_{rs}+\eta_{\mu s}\eta_{vr}\right)\left(\frac{\rho}{4m^2}\right)^3 F\left(1,1,\frac{9}{2};-\frac{\rho}{4m^2}\right)\\
&-i\frac{\pi^2 m^2\left(k^2-m^2\right)}{12\left(4m^2+k^2-i0\right)}\left(\eta_{\mu s}k_r k_v+\eta_{\mu v}k_r k_s+\eta_{rs}k_\mu k_v+\eta_{rv}k_\mu k_s\right)\\
&\times\left[\frac{1}{2}\left(\ln m^2+\frac{1}{3}\right)+\frac{1}{5}\left(\ln m^2+\frac{5}{6}\right)\frac{k^2}{4m^2}\right]\\
&-i\frac{\pi^2 m^2}{8\left(4m^2+k^2-i0\right)}\left(\eta_{\mu s}k_r k_v+\eta_{\mu v}k_r k_s+\eta_{rs}k_\mu k_v+\eta_{rv}k_\mu k_s\right)\\
&\times m^2\left[\left(\ln m^2+\frac{2}{3}\right)+\frac{k^2}{12}+\frac{k^2}{30}\frac{k^2}{4m^2}\right]-\frac{i\pi^2 m^4}{40\left(4m^2+k^2-i0\right)}\frac{k^2}{4m^2}\\
&-i\frac{\pi^2 m^2}{10}\left(\eta_{\mu s}k_r k_v+\eta_{\mu v}k_r k_s+\eta_{rs}k_\mu k_v+\eta_{rv}k_\mu k_s\right)\\
&\times\frac{k^2-m^2}{21\left(4m^2+k^2-i0\right)}F\left(1,1,\frac{9}{2};-\frac{\rho}{4m^2}\right)\left(\frac{k^2}{4m^2}\right)^2\\
&-i\frac{\pi^2}{12}k_\mu k_r k_v k_s\left[\left(\ln m^2+\frac{3}{4}\right)+\frac{k^2-4m^2}{2\left(4m^2+k^2-i0\right)}\right]\\
&-i\frac{\pi^2 m^2}{30}k_\mu k_r k_v k_s\frac{k^2-m^2}{4m^2+k^2-i0}\frac{k^2}{4m^2}F\left(1,1,\frac{7}{2};-\frac{k^2}{4m^2}\right)+\sum_{n=0}^{\infty}a_n\lambda^n.
\end{aligned}
\tag{8.7}
$$

Again, the exact result for our four-dimensional convolution becomes

$$
\begin{aligned}
&\Sigma_{GM\mu\nu rs}(k)=k_\mu k_r\left(\rho+m^2-i0\right)^{-1}*k_v k_s\left(\rho+m^2-i0\right)^{-1}\\
&=i\frac{m^2\pi^2}{2}\left(\eta_{\mu r}k_v k_s+\eta_{vs}k_\mu k_r\right)\times\left[\frac{1}{3}\left(\ln m^2+\frac{1}{12}\right)+\frac{1}{5}\frac{\rho}{4m^2}\left(\ln m^2+\frac{13}{15}\right)\right]\\
&+i\frac{m^2\pi^2}{30}\left(\eta_{\mu r}k_v k_s+\eta_{vs}k_\mu k_r\right)\frac{\rho}{4m^2}\\
&\times\left[F\left(1,1,\frac{7}{2};-\frac{\rho}{4m^2}\right)+\frac{1}{7}F\left(1,1,\frac{9}{2};-\frac{\rho}{4m^2}\right)\right]\\
&-i\frac{\pi^2 m^4}{4}\left(\eta_{\mu r}\eta_{vs}+\eta_{\mu v}\eta_{rs}+\eta_{\mu s}\eta_{vr}\right)\times\left\{\left[\frac{1}{2}-\frac{2}{3}\frac{\rho}{4m^2}-\frac{8}{15}\left(\frac{\rho}{4m^2}\right)^2\right]\right.\\
&\left.\times\left(\ln m^2+1\right)-\frac{1}{2}\left[\frac{3}{2}-\frac{1}{9}\left(\frac{\rho}{4m^2}\right)+\frac{52}{225}\left(\frac{\rho}{4m^2}\right)^2\right]\right\}\\
&-i\frac{2\pi^2 m^4}{105}\left(\eta_{\mu r}\eta_{vs}+\eta_{\mu v}\eta_{rs}+\eta_{\mu s}\eta_{vr}\right)\left(\frac{\rho}{4m^2}\right)^3 F\left(1,1,\frac{9}{2};-\frac{\rho}{4m^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&-i\frac{\pi^2 m^2\left(k^2-m^2\right)}{12\left(4m^2+k^2-i0\right)}\left(\eta_{\mu s}k_r k_v+\eta_{\mu v}k_r k_s+\eta_{rs}k_\mu k_v+\eta_{rv}k_\mu k_s\right)\\
&\times\left[\frac{1}{2}\left(\ln m^2+\frac{1}{3}\right)+\frac{1}{5}\left(\ln m^2+\frac{5}{6}\right)\frac{k^2}{4m^2}\right]\\
&-i\frac{\pi^2 m^2}{8\left(4m^2+k^2-i0\right)}\left(\eta_{\mu s}k_r k_v+\eta_{\mu v}k_r k_s+\eta_{rs}k_\mu k_v+\eta_{rv}k_\mu k_s\right)\\
&\times m^2\left[\left(\ln m^2+\frac{2}{3}\right)+\frac{k^2}{12}+\frac{k^2}{30}\frac{k^2}{4m^2}\right]-\frac{i\pi^2 m^4}{40\left(4m^2+k^2-i0\right)}\frac{k^2}{4m^2}\\
&-i\frac{\pi^2 m^2}{10}\left(\eta_{\mu s}k_r k_v+\eta_{\mu v}k_r k_s+\eta_{rs}k_\mu k_v+\eta_{rv}k_\mu k_s\right)\\
&\times\frac{k^2-m^2}{21\left(4m^2+k^2-i0\right)}F\left(1,1,\frac{9}{2};-\frac{\rho}{4m^2}\right)\left(\frac{k^2}{4m^2}\right)^2\\
&-i\frac{\pi^2}{12}k_\mu k_r k_v k_s\left[\left(\ln m^2+\frac{3}{4}\right)+\frac{k^2-4m^2}{2\left(4m^2+k^2-i0\right)}\right]\\
&-i\frac{\pi^2 m^2}{30}k_\mu k_r k_v k_s\frac{k^2-m^2}{4m^2+k^2-i0}\frac{k^2}{4m^2}F\left(1,1,\frac{7}{2};-\frac{k^2}{4m^2}\right)
\end{aligned}
\tag{8.8}
$$

We have to deal with 9 diagrams of this kind.

Accordingly, our desired self-energy total is a combination of $\Sigma_{G\alpha_1\alpha_2\alpha_3\alpha_4}(k)$ and $\Sigma_{GM\alpha_1\alpha_2\alpha_3\alpha_4}(k)$.

9. Self Energy of the Axion

Here a typical term of the self-energy is:

$$\Sigma_{vr}(k)=k_v k_r\left(\rho+m^2-i0\right)^{-1}*\left(\rho-i0\right)^{-1}. \tag{9.1}$$

In four dimensions one has

$$k_v k_r\left(\rho+m^2-i0\right)^{-1}*\left(\rho-i0\right)^{-1}=\int\frac{k_v k_r}{\left(k^2+m^2-i0\right)\left[\left(p-k\right)^2-i0\right]}\mathrm{d}^4k. \tag{9.2}$$

with the Feynman parameters used above we obtain

$$
\begin{aligned}
&k_v k_r\left(\rho+m^2-i0\right)^{\lambda-1}*\left(\rho-i0\right)^{\lambda-1}\\
&=i\int_0^1 x^{-\lambda}\left(1-x\right)^{-\lambda}\int\frac{k_v k_r}{\left[\left(k-px\right)^2+a\right]^{2-\lambda}}\mathrm{d}^4k\mathrm{d}x,
\end{aligned}
\tag{9.3}
$$

where

$$a=\left(p^2+m^2\right)x-p^2x^2. \tag{9.4}$$

We evaluate the integral (9.3) and find

$$
\begin{aligned}
&k_v k_r\left(\rho+m^2-i0\right)^{\lambda-1}*\left(\rho-i0\right)^{\lambda-1}\\
&=i\frac{\eta_{vr}m^{2+4\lambda}\pi^2}{4}\frac{\Gamma\left(2+\lambda\right)}{\Gamma\left(1-\lambda\right)}\Gamma\left(-1-2\lambda\right)F\left(-1-2\lambda,1-\lambda,3;-\frac{\rho}{m^2}\right)
\end{aligned}
$$

$$+\frac{ik_v k_r m^{4\lambda}\pi^2}{6}\frac{\Gamma(3+\lambda)}{\Gamma(1-\lambda)}\Gamma(-2\lambda)F\left(-2\lambda,1-\lambda,4;-\frac{\rho}{m^2}\right). \tag{9.5}$$

Self-Energy Evaluation for $\lambda = 0$

Once again, we Laurent-expand, this time (9.5) around $\lambda = 0$, encountering

$$\begin{aligned} & k_v k_r\left(\rho + m^2 - i0\right)^{\lambda-1} * \left(\rho - i0\right)^{\lambda-1} \\ & = i\pi^2\left\{\frac{1}{2\lambda}\left(\frac{\eta_{vr}m^2}{4}-\frac{1}{3}k_v k_r\right)+\frac{\eta_{vr}m^2}{4}\left[\left(1+\frac{1}{3}\frac{\rho}{m^2}\right)\left(\ln m^2+\frac{1}{2}\right)\right.\right. \\ & \left.-\left(1+\frac{1}{6}\frac{\rho}{m^2}\right)\right]-\frac{k_v k_r}{3}\left(\ln m^2+\frac{3}{4}\right) \\ & \left.+\frac{1}{4}\left(\frac{\rho}{m^2}\right)\left[\frac{\eta_{vr}m^2}{12}\frac{\rho}{m^2}-\frac{k_v k_r}{3}\right]F\left(1,1,5;-\frac{\rho}{m^2}\right)+\sum_{n=1}^{\infty}a_n\lambda^n\right\} \end{aligned} \tag{9.6}$$

The λ-independent term gives the exact convolution result we are looking for:

$$\begin{aligned} \Sigma_{vr}(k) &= k_v k_r\left(\rho + m^2 - i0\right)^{-1} * \left(\rho - i0\right)^{-1} \\ &= i\pi^2\left\{\frac{\eta_{vr}m^2}{4}\left[\left(1+\frac{1}{3}\frac{\rho}{m^2}\right)\left(\ln m^2+\frac{1}{2}\right)\right]\right. \\ &\left.-\left(1+\frac{1}{6}\frac{\rho}{m^2}\right)\right]-\frac{k_v k_r}{3}\left(\ln m^2+\frac{3}{4}\right) \\ &\left.+\frac{1}{4}\left(\frac{\rho}{m^2}\right)\left[\frac{\eta_{vr}m^2}{12}\frac{\rho}{m^2}-\frac{k_v k_r}{3}\right]F\left(1,1,5;-\frac{\rho}{m^2}\right)\right\} \end{aligned} \tag{9.7}$$

10. Discussion

We have developed above a quantum field theory (QFT) of Eintein's gravity (EG), that is both unitary and finite, by appealing to the Schwinger-Feyman variational principle. We emphatically avoid the functional integral method. Our results critically depend on the use of a rather novel constraint the we introduced in defining the EG-Lagrangian. Laurent expansions were also an indispensable tool for us. As sgtated, in order to quantify the theory we appealed to the variational principle of Schwinger-Feynman's. This process leads to just one graviton type ϕ^{12}. The underlying mathematics used in this effort has been developed by Bollini *et al.* [1] [2] [3] [4] [5]. This mathematics is powerful enough so as to be able to quantize non-renormalizable field theories [1] [2] [3] [4] [5]. We have evaluated here in finite and exact fashion, for the first time as far as we know, several quantities:

- the graviton's self-energy in the EG-field. This requires full use of the theory of distributions, appealing to the possibility of creating with them a ring with divisors of zero.
- the above self-energy in the added presence of a massive scalar field (axions, for instance). Two types of diagram ensue: the original ones of the pure EG field plus the ones originated by the addition of a scalar field.

- the axion's self-energy.
- Our central results revolve around Equation (6.6), Equation (8.8), and Equation (9.7), corresponding to the graviton's self-energy, without and with the added presence of axions. Also, we give the axion's self-energy.

As a final remark, we would like to point out that our formula for convolutions is a mathematical definition and not a regularization.

Conflicts of Interest

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

References

[1] Bollini, C.G., Escobar, T. and Rocca, M.C. (1999) *International Journal of Theoretical Physics*, **38**, 2315. https://doi.org/10.1023/A:1026623718239

[2] Bollini, C.G. and Rocca, M.C. (2004) *International Journal of Theoretical Physics*, **43**, 1019. https://doi.org/10.1023/B:IJTP.0000048599.21501.93

[3] Bollini, C.G. and Rocca, M.C. (2004) *International Journal of Theoretical Physics*, **43**, 59. https://doi.org/10.1023/B:IJTP.0000028850.35090.24

[4] Bollini, C.G., Marchiano, P. and Rocca, M.C. (2007) *International Journal of Theoretical Physics*, **46**, 3030. https://doi.org/10.1007/s10773-007-9418-y

[5] Plastino, A. and Rocca, M.C. (2018) *Journal of Physics Communications*, **2**, Article ID: 115029. https://doi.org/10.1088/2399-6528/aaf186

[6] Plastino, A. and Rocca, M.C. (2019) Gupta-Feynman Based Quantum Field Theory of Einstein's Gravity. https://www.researchgate.net/publication/336406184_Gupta-Feynman_based_Quantum_Field_Theory_of_Einstein's_Gravity

[7] Gupta, S.N. (1968) *Proc. Pys. Soc. A*, **65**, 161.

[8] Grothendieck, A. (1955) *Memoirs of the American Mathematical Society*, **16**. https://doi.org/10.1090/memo/0016

[9] Sebastiao e Silva, J. (1958) *Mathematische Annalen*, **136**, 38. https://doi.org/10.1007/BF01350287

[10] Feynman, R.P. (1963) *Acta Physica Polonica*, **24**, 697.

[11] Visconti, A. (1969) Quantum Field Theory. Pergamon Press, Oxford.

[12] Delbourgo, R. and Prasad, V.B.J. (1975) *Journal of Physics G: Nuclear Physics*, **1**, 377. https://doi.org/10.1088/0305-4616/1/4/001

[13] Barci, D.G., Bollini, C.G. and Rocca, M.C. (1995) *Il Nuovo Cimento*, **108**, 797. https://doi.org/10.1007/BF02731021

[14] Kleinert, H. (2016) Particles and Quantum Fields. Free Web Version. https://pt.b-ok2.org/book/2747268/e795bc https://doi.org/10.1142/9915

[15] Peccei, R.D. (2008) The Strong CP Problem and Axions. In: Kuster, M., Raffelt, G. and Beltrn, B., Eds., *Axions: Theory, Cosmology, and Experimental Searches*, Lecture Notes in Physics, Vol. 741, Springer, Heidelberg, 3-17. https://doi.org/10.1007/978-3-540-73518-2_1

Journal of Modern Physics, 2020, 11, 395-406
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

How Massive Are the Superfluid Cores in the Crab and Vela Pulsars and Why Their Glitch-Events Are Accompanied with under and Overshootings?

A. A. Hujeirat[1], R. Samtaney[2]

[1]IWR, Universität Heidelberg, Heidelberg, Germany
[2]Applied Mathematics and Computational Science, CEMSE Division, KAUST, Thuwal, KSA
Email: AHujeirat@uni-hd.de, ravi.samtaney@kaust.edu.sa

How to cite this paper: Hujeirat, A. A. and Samtaney, R. (2020) How Massive Are the Superfluid Cores in the Crab and Vela Pulsars and Why Their Glitch-Events Are Accompanied with under and Overshootings? *Journal of Modern Physics*, **11**, 395-406.
https://doi.org/10.4236/jmp.2020.113025

Received: February 17, 2020
Accepted: March 13, 2020
Published: March 16, 2020

Abstract

The Crab and Vela are well-studied glitching pulsars and the data obtained so far should enable us to test the reliability of models of their internal structures. Very recently it was proposed that glitching pulsars are embedded in bimetric spacetime: their incompressible superfluid cores (SuSu-cores) are embedded in flat spacetime, whereas the ambient compressible and dissipative media are enclosed in Schwarzschild spacetime. In this letter we apply this model to the Crab and Vela pulsars and show that a newly born pulsar initially of $1.25M_{\odot}$ and an embryonic SuSu-core of $0.029M_{\odot}$ could evolve into a Crab-like pulsar after 1000 years and into a Vela-like pulsar 10,000 years later to finally fade away as an invisible dark energy object after roughly 10 Myr. Based thereon we infer that the Crab and the Vela pulsars should have SuSu-cores of $0.15M_{\odot}$ and $0.55M_{\odot}$, respectively. Furthermore, the under- and overshootings phenomena observed to accompany the glitch events of the Vela pulsar are rather a common phenomenon of glitching pulsars that can be well-explained within the framework of bimetric spacetime.

Keywords

Relativity: Numerical, General, Black Hole Physics, Magnetars, Neutron Stars, Pulsars, Superfluidity, Superconductivity, Gluons, Quarks, Quantum Chromodynamics (QCD)

1. Observational Constraints and Methodology

The Crab and Vela pulsars are well-known and extensively studied pulsars (see

[1]-[7] and the references therein). In **Table 1** we summarize their basic observational data relevant for the present discussion. The SuSu-Scenario relies on solving the TOV equation in combination with the equations of torque balance between the incompressible superfluid cores, whose dynamics obey the Onsager-Feymann equation, and an overlying shell of compressible and dissipative matter (see Sec. 2 and Eq. 10 in [8]).

In [9] these equations were solved at the background of a bimetric space-time (see **Figure 1**). Unlike the original model [8], in which the spin-down of the SuSu-core is set to follow a priori given sequence of values $\left\{\Omega_c^n\right\}$, in the present work however, the SuSu-core is set to undergo an abrupt spin-down, if the difference between its Eigen rotation and that of the ambient medium surpasses a critical value $\left\{\Delta\Omega_{cr}^n\right\}$, *i.e.* if $\Delta\Omega_{c-am}^n = \Omega_c^n - \Omega_{am}^n \geq \Delta\Omega_{cr}^n$, where "*n*" and "*am*" refer to the order of the elements in the relevant sequence and to the ambient medium, respectively. This approach is more consistent than the former, as the elements of $\Delta\Omega_{cr}^n$ are determined here through the rate of loss of rotational energy of the entire star and that these should overlap the current values observed in the Crab and Vela pulsars.

The strategy of obtaining the optimal values here relies on using a global iterative solution procedure that takes the following constraints into account (see also **Table 1**):

1) The elements of the sequence $\left\{\frac{\Delta\Omega_g}{\Omega}\right\}^n$ must fulfill the three conditions:

- $\left\{\frac{\Delta\Omega_g}{\Omega}\right\}^n \xrightarrow[n\to 0]{} 0$, which means that the media in both the core and in the surrounding shell must have identical rotational frequency initially.
- $\left.\frac{\Delta\Omega_g}{\Omega}\right|_{n=N_0} = \left.\frac{\Delta\Omega_g}{\Omega}\right|_{Crab} = 4\times 10^{-9}$ and $\left.\frac{\Delta\Omega_g}{\Omega}\right|_{n=N_1} = \left.\frac{\Delta\Omega_g}{\Omega}\right|_{Vela} = 2.33\times 10^{-6}$, *i.e.* the elements number N_0 and $N_1 (\gg N_0)$ of the sequence $\left\{\frac{\Delta\Omega_g}{\Omega}\right\}^n$ must be identical to the observed values of the Crab and to the Vela pulsars, respectively.

Table 1. A list of the main observational data of the Crab and Vela pulsars relevant for the present study (see [2]-[7] and the references therein).

	Crab	Vela
Mass ($M_\odot$)	1.4	1.8
Age (kyr)	1.24	11.3
$B(10^{12}\,\mathrm{G})$	4.875	4.35
$\Omega\left(\mathrm{s}^{-1}\right)$	200	70
$\Delta\Omega_g/\Omega$	4×10^{-9}	2.338×10^{-6}
$\Delta t_g\,(\mathrm{yr})$	1.6	2.5

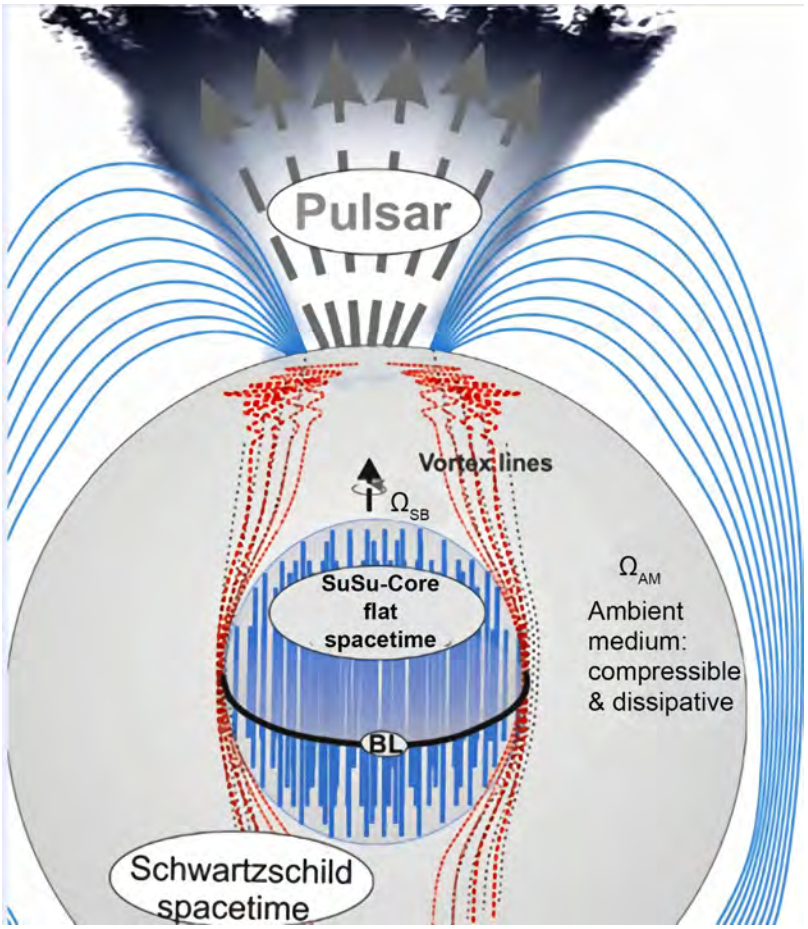


Figure 1. A schematic description of the bimetric spacetime inside glitching pulsars: The incompressible superfluid core is embedded in a Minkowski spacetime whereas the ambient media are enclosed in Schwarzschild spacetime.

- $\left\{\frac{\Delta\Omega_g}{\Omega}\right\}^n \xrightarrow[n\to\infty]{} \alpha_\infty < \infty$, *i.e.* the sequence must converge to a finite value. Moreover, our test calculations have shown that for $t \gg t_{Vela}$, $\frac{\partial}{\partial t}\left(\left\{\frac{\Delta\Omega_g}{\Omega}\right\}^n\right) < 0$ as otherwise the magnetic field would fail to spin-down the crust and therefore to surpass $\Delta\Omega_{cr}$ required for triggering a prompt spin-down of the core into the next lower energy state.

Indeed, one possible sequence which fulfills the above-mentioned constraints, though it might not be unique, is shown in Figure 2.

2) The initial conditions used here are $\Omega_0 = \Omega(t=0) = 1440$ Hz (see [10] and the references therein). Here both the core and the ambient medium are set to initially rotate with the same frequency, *i.e.* $\Omega_c^n = \Omega_{am}^n(t=0) = \Omega_0$. The initial total mass of the pulsars and that of the core are taken to be: $M_c = \epsilon_0 M_\odot$ and $M_0 = M(t=0) = \epsilon_1 M_\odot$, respectively, where ϵ_0 and ϵ_1 are parameters whose values are determined through a global iteration procedure. The initial magnetic field strength is taken to be $B(t=0) = 10^{13}$ Gauss.

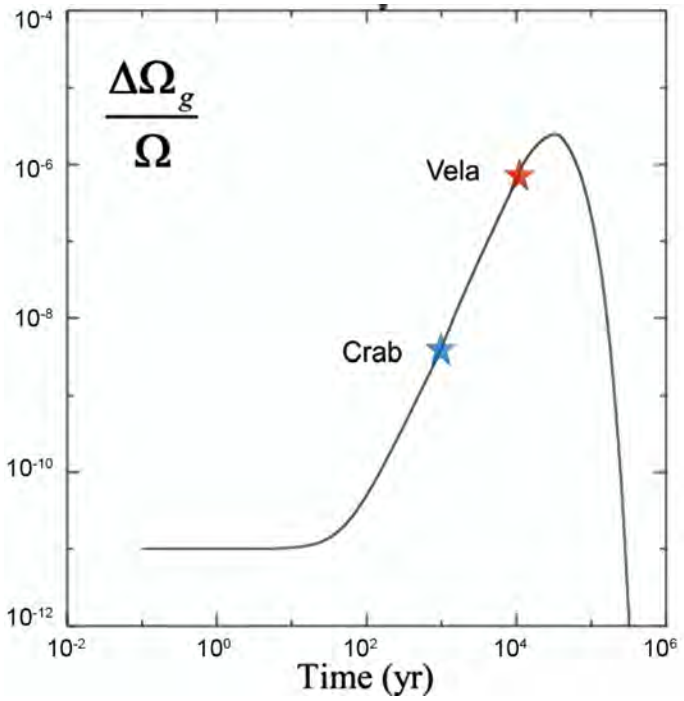


Figure 2. The elements of the quantum sequence $\Delta\Omega_g^n/\Omega$ are shown as function of time in year units. Both the core and the ambient medium are set to rotate with the same frequency 1440s^{-1} initially. As the pulsar cools down, $\Delta\Omega_g^n/\Omega$ starts increasing to reach 4 $\times 10^{-9}$ after approximately 1000 yr, (which corresponds to the Crab phase/blue-star) and 8.15 $\times 10^{-7}$ after 11,000 yr (which corresponds to Vela phase/red-star). $\Delta\Omega_g^n/\Omega$ here is measured in units of $\Omega = 200/\mathrm{s}$.

3) The elements of the sequence $\{\Omega_{am}\}^n$ are obtained through the energy balance equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\Omega_{am}^2}\right) = -\alpha_{EM}\frac{B^2}{I_{am}}, \tag{1}$$

where I_{am} is the inertia of the ambient compressible dissipative medium, which, due to the increase of the SuSu-core, must decrease on the cosmic time and $\alpha_{EM} = 4.9\times 10^{-4}$ is a non-dimensional constant.

Thus, the time-evolution of the core's rotational frequency proceeds as follows: for a given Ω_c^n, the ambient medium is set to decrease its frequency continuously with time through the emission magnetic dipole radiation. This implies that the difference $\Delta\Omega_{c-am}$ should increase with time until $\Delta\Omega_{c-am}$ has surpassed the critical value $\geq \Delta\Omega_{cr}$, In this case three events are expected to occur promptly:

- The rigid-body rotating core changes abruptly its rotational state from $E_c^{n,rot} = \frac{1}{2}I_c^n\left(\Omega_c^n\right)^2$ into the next the quantum-mechanically permitted lower energy state: $E_c^{n+1,rot} = \frac{1}{2}I_c^{n+1}\left(\Omega_c^{n+1}\right)^2$. This process is associated with ejection of a certain number of vortices into the boundary layer (BL) between the core

and the overlying dissipative medium.

- The ejected vortices by the core are then absorbed by the differentially rotating dissipative medium and re-distributed viscously. Hence the medium in the BL would experience the prompt spin-up: $\Omega_{am}^n \to \Omega_{am}^n + \Delta\Omega_g^n$, where $\Delta\Omega_g^n$ is deduced from the sequence $\left\{\frac{\Delta\Omega_g}{\Omega}\right\}^n$.
- The radius of the core is set to increases as dictated by the Onsager-feynmann equation:

$$\oint V \cdot \mathrm{d}\ell = \frac{2\pi\hbar}{m} N, \tag{2}$$

where $V, \ell, \hbar, m, N$ denote the velocity vector, the vector of line-element, the reduced Planck constant, mass of the superfluid particle pair and the number of vortices, respectively (see [11] for further details). Imposing zero-torque condition on the incompressible SuSu-core, *i.e.*, $\frac{\mathrm{d}}{\mathrm{d}t}(\Omega I)_c = 0$. We then obtain the following recursive relation:

$$\begin{aligned} &(\Omega S)_c^{n+1} = (\Omega S)_c^n \\ &\Rightarrow S^{n+1} = \left(\frac{\Omega_c^n}{\Omega_c^{n+1}}\right) S^n \end{aligned} \tag{3}$$

where $S_c^n \doteq \pi\left(R_c^n\right)^2$ and R_c^n correspond to the cross-sectional area of the SuSu-core and to the corresponding radius, respectively. The increase in the dimension of the core implies that the matter in the geometrically thin boundary layer between the SuSu-core and the ambient medium should undergo a crossover phase transition into an incompressible superfluid, whose total energy density saturates around the critical value $\rho_{cr} \approx 6\rho_0$ (see [12] and the references therin). The growth of the core proceeds on the cosmic time scale and ends when the pulsar has metamorphosed entirely into a maximally compact invisible dark energy object and therefore becomes observationally indistinguishable from a stellar black hole.

2. Solution Procedure & Results

The set of equations consists of the TOV equation for modeling the compressible dissipative matter in the shell overlaying the incompressible gluon-quark superfluid core, whereas the latter is set to obey the zero-torque condition and to dynamically evolve according to the Onsager-Feymann equation (for further details see Sec. 2 and Eq. 10 in [8]).

The global iteration loop is designed here to find the optimal values of the parameters: α_0, α_1, the elements of the sequence $\Delta\Omega_{cr}^n$ and the decay rate of the magnetic field. These values should fulfill the initial and final conditions, the currently observed values of the time passages between two successive glitch-events Δt_g both of the Crab and the Vela pulsars, the current observed values of their magnetic fields masses.

Indeed, our intensive computations reveal that optimal fitting may be achieved for $M_c(t=0) \approx 0.029 M_\odot$, a sequence of $\Delta\Omega_g^n$, whose elements are shown in Figure 2. In Figure 3, Figure 4 the optimal values of $\Delta\Omega_{cr}^n$ and $\Delta\Omega_g^n$ are shown versus cosmic time, whereas in Figure 5 we show the time-development of the rotational frequencies of the core, the ambient medium and of the magnetic field during the first 10 to 100 years after the birth of the pulsar. The long-term evolution of the magnetic field and the growing mass of the core and of the entire object are shown in Figure 6 and Figure 7. Here the mass of the pulsar's core grows with time to reach $0.15 M_\odot$ after 1000 years and reaches $0.55 M_\odot$ after 11,000 years; hence reproducing the exact total masses of $1.4 M_\odot$ for the Crab and $1.8 M_\odot$ for the Vela pulsars as revealed from observations. The relative ratio of inertia of both cores reads $I^{Crab}/I^{Vela} \approx 3/20$.

Due to the incompressible, superfluid and supreconducting character of the core, the evolution of the magnetic field is solely connected to the dynamics of the ambient compressible and dissipative matter in the shell as well as to its dimensions (see [13] for further details on the physical aspects of compressibility of fluid flows). As the mass and dimension of the core grow with time, the surrounding shell must shrink. In this case, conservation of the magnetic flux should strengthen the magnetic field intensity. This interplay between the loss of magnetic energy due to loss of rotational energy and enhancement by conservation of magnetic flux in combination with dynamo action and other mechanisms,

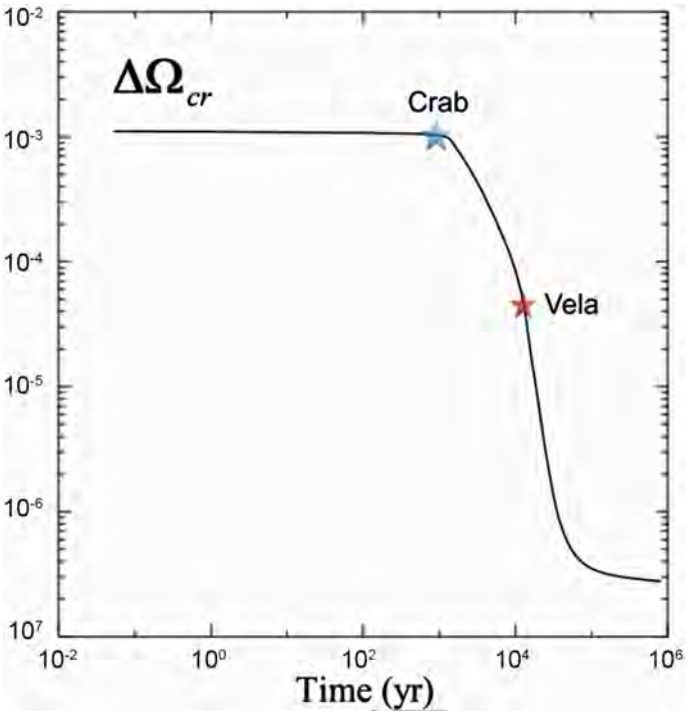


Figure 3. The elements of the sequence $\Delta\Omega_{cr}^n$ versus cosmic time. Each element corresponds to the critical difference between the rotational of frequency of the core and that of the ambient medium, beyond which the core undergoes a prompt spin-down to the next lower energy state.

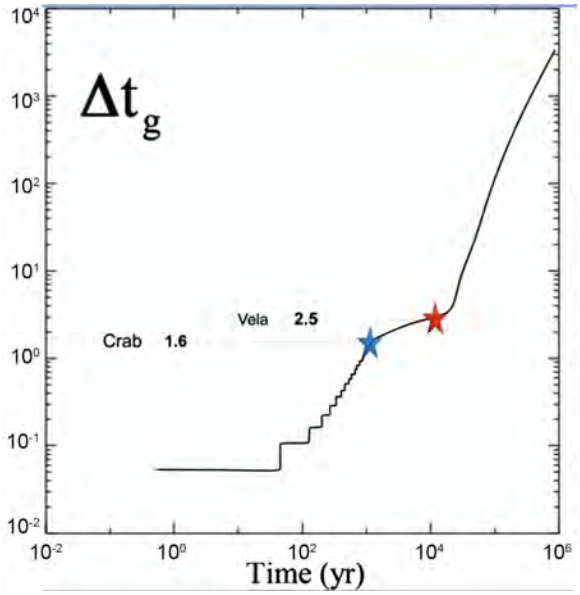


Figure 4. The elements of the sequence Δt_g versus cosmic time. Each element corresponds to the time passage between two successive glitch events. The actual values that correspond to the Crab and Vela pulsars are shown in blue and red stars. These time passages increase dramatically with time due to the decay of both magnetic field intensity and rotational energy.

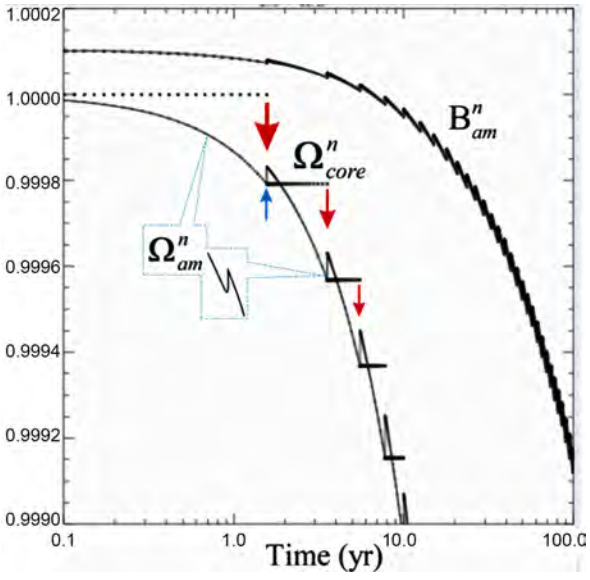


Figure 5. The elements of the sequences $\Omega_c^n, \Omega_{am}^n$ and B_{am}^n that corresponds to the rotational frequencies of the core, the ambient medium and of the magnetic field during the first 10 years. During a passage of time between two successive glitches, the core rotates rigidly with a constant frequency $\Omega_c^n = \Omega_{core}^n$ (dotted line), whereas the ambient medium spin-down in a continuous manner (solid line). During the glitch, the core spin-down abruptly, triggering a prompt spin-up of the ambient medium in the boundary layer between the rigid-body rotating core and the overlying differentially rotating medium. The enhanced spin-up of the ambient medium in combination with the decreasing volume enclosing this matter gives rise to magnetic field B_{am} which evolves in a similar discrete manner (dash-dot).

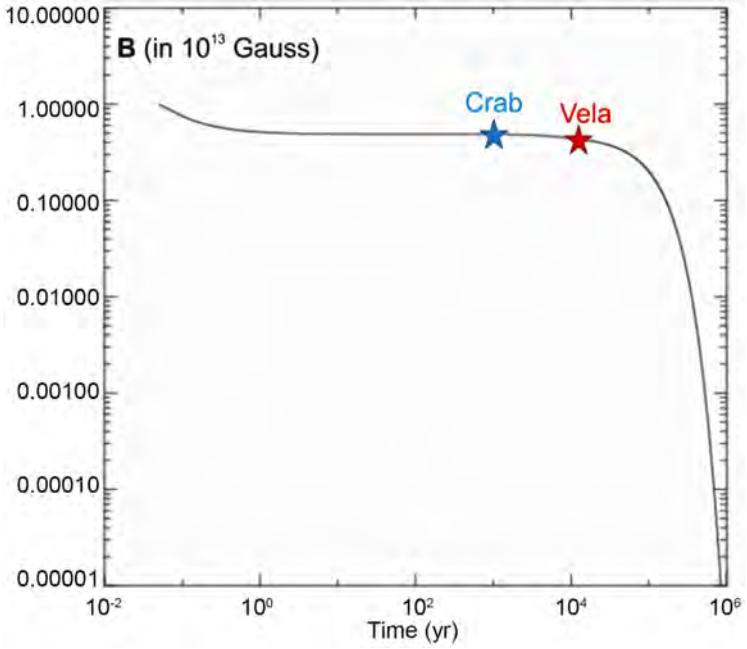


Figure 6. The cosmic evolution of the magnetic field-B, of a newly born pulsar in units of 10^{13} G. The superimposed blue and red stars correspond to the current B of the Crab and Vela pulsars.

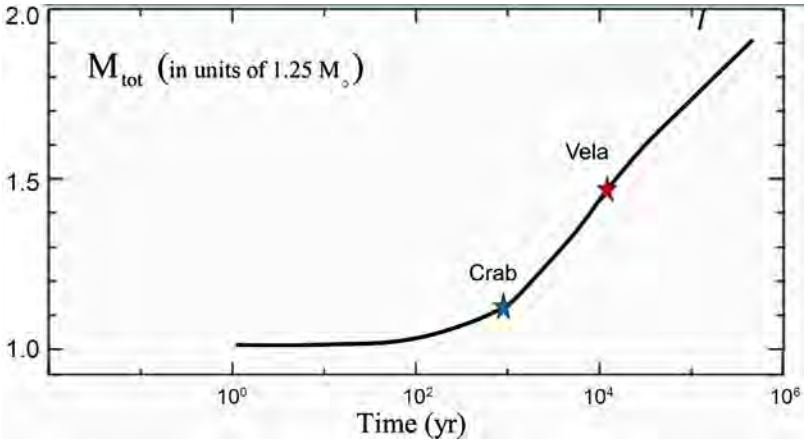


Figure 7. The mass growth of a newly born pulsar having initially $M_0 = 1.25M_\odot$ and an embryonic SuSu-core of $0.029M_\odot$. After approximately 1000 yr the pulsar recover the mass of the Crab ($M_0 = 1.4M_\odot$) and 10,000 yr later $M_0 = 1.8M_\odot$ that corresponds to the mass of the Vela. At the end of the luminous life time, which lasts for approximately 10 million years, the pulsar enters the dark phase with a total mass of $2.5M_\odot$, which corresponds to a maximally compact invisible dark energy object. The plotted mass here is in units of M_0.

may clarify the very weak decay of magnetic field as pulsars evolve from the Crab to the Vela phase. Mathematically, let the magnetic energy in a shell of a newly born pulsar be:

$$E_M = \int \frac{B^2}{8\pi} \mathrm{d}v \sim \frac{B^2}{6}\left(R_\star^3 - R_c^3\right), \tag{4}$$

where $R_\star$ denotes the pulsar's radius. Assuming E_M to roughly decay as the rotational energy E_Ω, then we obtain:

$$B^- = \alpha_\Omega \Omega_{am}^2 M_{am}^{1/2} \sqrt{\frac{R_\star^2 - R_c^2}{R_\star^3 - R_c^3}}, \tag{5}$$

where α_Ω is constant coefficient.

On the other hand, dynamo action in combination with magnetic flux conservation and other enhancement mechanisms would contribute positively to the magnetic field, that can, for simplicity absorbed in the term: $B^+ = \alpha_B \Big/ \left(R_\star^2 - R_c^2\right)$. The coefficient α_B is set to ensure that the magnetic field remains in the very sub-equipartition regime. Hence the interplay between magnetic loss and enhancement would yield an effective magnetic field that evolves according to:

$$B_{tot} = \frac{\alpha_B}{R_\star^2 - R_c^2} - \alpha_\Omega \Omega_{am}^2 M_{am}^{1/2} \sqrt{\frac{R_\star^2 - R_c^2}{R_\star^3 - R_c^3}}. \tag{6}$$

Consequently, our model predicts that the decreasing volume of the shell enclosing the ambient medium in combination with dynamo action in the boundary layer could potentially be the mechanism that keeps the decay of magnetic fields in pulsars extremely weak.

In fact our model predicts the glitch activity of a newly born pulsar, which evolves into a Crab phase, followed by a Vela phase and finally by an invisible phase, to be approximately two orders of magnitude larger than it was estimated by other models (see Figure 8 to be compared to [1] [5]). According to our

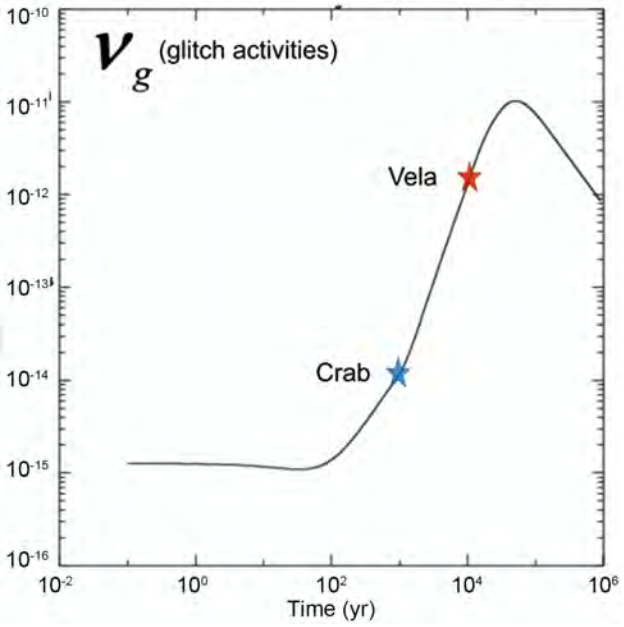


Figure 8. The glitch activity of a newly born pulsar versus cosmic time. In the very early times, the pulsar underwent millions of glitches, though the total ejected rotational energy was relatively very low. These activities start to be significant as the pulsar ages and become maximally effective between 100 and 60,000 years, followed by a decreasing phase, during which time-passages between successive glitches become increasingly longer.

model pulsar may undergo millions or up to billions of glitches during their luminous life time with passages of time between two successive glitch events that range from nanoseconds in the very early time up to hundreds or even thousands of years toward the end of their luminous life times (see **Figure 4**). The vast difference in the evolution of glitch activity between the two approaches here may be attributed to the strong non-uniformity of time-duration between glitch events.

Moreover, the model also predicts the occurrence of under- and overshootings that have been observed to accompany the glitch events in the Vela pulsar (see [14] and the references therein). In the case of the Vela, when the core expels certain number of vortices and moves to the next lower energy state, the enhanced rotational energy of the matter in the BL amplifies the magnetic field strength. Due to the non-locality of magnetic fields[1], this enhancement is communicated to the crust via Alfven waves, V_A, whereas the excess of rotational energy is communicated via shear viscosity with an effective propagational velocity V_{vis}. As these two speeds are generally different with $V_A > V_{vis}$ in most cases, the time-delay in the arrival of communication enforces the crust to react differently. Specifically, the arrival of magnetic enhancement prior to the rotational one leaves the crust subject to an enhanced magnetic braking and therefore to a stronger reduction of its rotational frequency (see the top panel of **Figure 9**).

Indeed, in the case of Vela, the propagational speed of Alfven waves may be esitmated to be of order $V_A \sim B/\sqrt{\rho} \approx 10^8\,\mathrm{cm/s}$. Hence the enhanced MFs in the BL would be communicate to the crust within $\delta\tau_{MF} = \Delta R/V_A = (R_\star - R_c)/V_A \approx 10^{-2}\,\mathrm{s}$. On the other hand, supplying the crust with rotational energy would proceed on the viscous time scale, which is estimated to be: $\delta\tau_{vis} = (\Delta R)^2/\nu_{vis}$ (see [15] and the references therein). Under normal astrophysical conditions we may safely assume that $V_{vis}/V_A = \alpha_2 \ll 1$ and that the length scale, ℓ_{vis}, over which viscous interaction occurs, ℓ_{vis} is much smaller than the width of the shell ΔR, or equevalently $\ell_{vis} = \alpha_2 \Delta R$, with $\alpha_2 = 1$. Therefore $\delta\tau_{Vis} = \Delta R/(\alpha_1\alpha_2 V_A) \approx 10^{-2}/(\alpha_1\alpha_2) \approx 1\,\mathrm{s}$, where we reasonably set $\alpha_1 = \alpha_2 = 0.1$. Consequently, the observed undershooting most likely results due to the time delay of the arrival of communications via magnetic fields and viscous torque, which amounts to $\delta\tau_{vis}/\delta\tau_{MF} = 1/(\alpha_1\alpha_2) \approx 100$. On the other hand, the observed overshooting can be attributed to the case in which the viscous front transporting rotational energy from the BL outwards has reached the crust. As $\delta\tau_{MF} \ll \delta\tau_{vis}$, at the end of $\delta\tau_{vis}$, the magnetic field intensity in the BL should have returned to values comparable or even lower than prior to the glitch event.

Moreover, the observed order in which undershooting followed by overshooting is an indication for a time-delay in the arrival of communication resulting from $V_A > V_{vis}$ and from the significant difference of the locations of the BL and the crust. This order is expected to reverse if $V_A < V_{vis}$.

I.e. In the absence of magnetic monopoles.

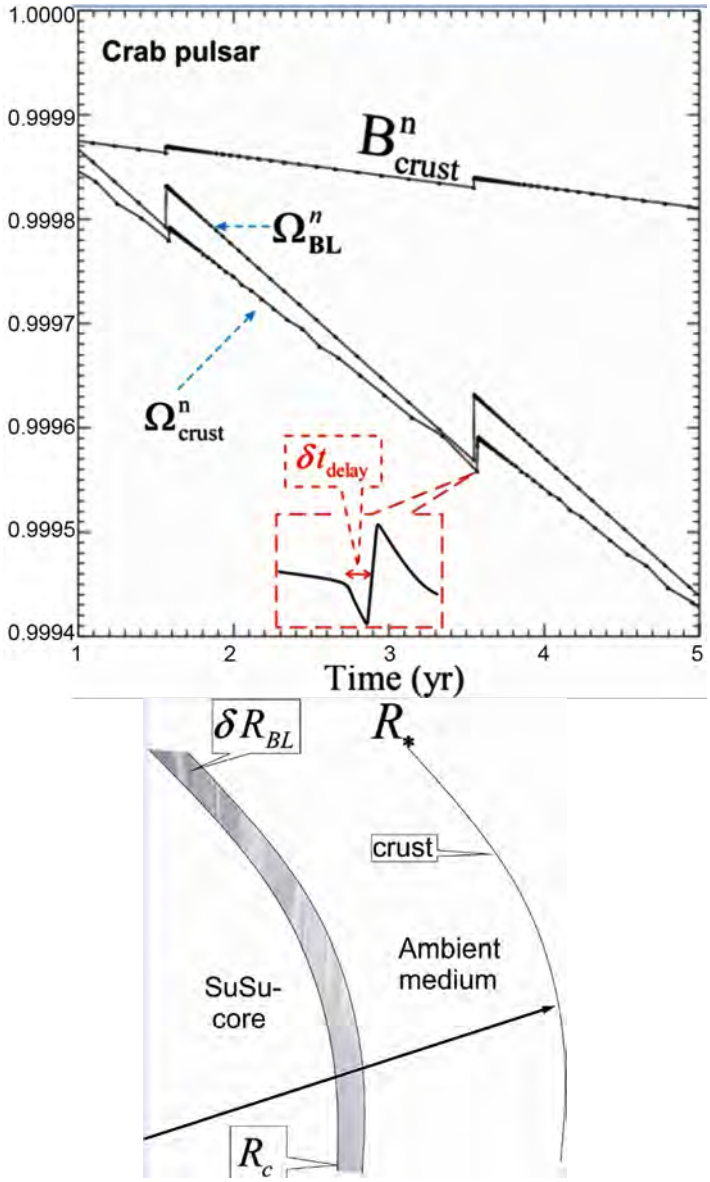


Figure 9. Under and overshooting of Ω_{am} shortly before and immediately after a glitch event (top Figure). At a given instant of time, *t*, the rotational frequency of the matter in the BL, Ω_{am}, differs from that of the crust Ω_{crust}. Due to their different locations (see lower panel), the response of the crust to the dynamical changes of matter in the BL depends strongly on the speed of communication via magnetic fields (e.g. Alfven waves) and shear viscosity, which, under most astrophysical conditions, are considered to be different.

In fact, the under- and overshooting here may indicate that MFs are insensitive to the momentary rotational frequency of the crust, but rather to the activity and dynamics of the matter in the BL.

Extending this analysis to both the Crab and Vela pulsars, the relative time-delays is expected to be: $\delta\tau_{vis}^{Crab}/\delta\tau_{vis}^{Vela} \sim \left(\Delta R^{Crab}/\Delta R^{Vela}\right)^2 \approx 3.4$ or equivalently, the undershooting in the case of the Crab is expected to last 3.4 sec compared to one second in the Vela case.

Finally, although the physics is entirely different, the situation here is strikingly similar to action of the solar dynamo, which is considered to be located in the so-called tachcline between the rigid-body rotating core and the overlying convection zone [16].

Acknowledgements

The calculations have been carried out using the computer cluster of the IWR, University of Heidelberg. RS acknowledges the use of KAUS baseline research funds

Conflicts of Interest

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

References

[1] Roy, J., Yashwant Gupta, Y. and Lewandowski, W. (2012) *MNRAS*, **424**, 2213-2221.

[2] Yu, M., Manchester, R.N., Hobbs, G., *et al.* (2013) *MNRAS*, **429**, 688-724.

[3] Yuan, J., Kou, F. and Wang, N. (2019) *AIP Conference Proceedings*, **2127**, Article ID: 020004.

[4] Eya, I.O. and Urama, J.O. (2014) *International Journal of Astrophysics and Space Science*, **2**, 16.

[5] Espinoza, C.M., Lyne, A.G., Stappers, B.W. and Kramer, C. (2011) *MNRAS*, **414**, 1679. https://doi.org/10.1111/j.1365-2966.2011.18503.x

[6] Fuentes, J.R., Espinoza, C.M., Reisenegger, A., *et al.* (2017) *A&A*, **608**, A131. https://doi.org/10.1051/0004-6361/201731519

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

[8] Hujeirat, A.A. (2018) *Journal of Modern Physics*, **9**, 4.

[9] Hujeirat, A.A. and Samtaney, R. (2019) *Journal of Modern Physics*, **9**, 4.

[10] Haensel, P., Lasota, J.P. and Zdunik, J.L. (1999) *A&A*, **344**, 151.

[11] Hujeirat, A.A. (2018) *Journal of Modern Physics*, **9**, 51-69. https://doi.org/10.4236/jmp.2018.91004

[12] Hujeirat, A.A. (2018) *Journal of Modern Physics*, **9**, 4.

[13] Hujeirat, A.A. and Thielemann, F.-K. (2009) *MNRAS*, **400**, 903. https://doi.org/10.1111/j.1365-2966.2009.15498.x

[14] Ashton, G., Lasky, P.D., Graber, V. and Palfreyman, J. (2019) Rotational Evolution of the Vela Pulsar during the 2016 Glitch.

[15] Hujeirat, A.A. and Thielemann, F.-K. (2009) *A&A*, **496**, 609. https://doi.org/10.1051/0004-6361/200810793

[16] Camenzind, M. (2007) Compact Objects in Astrophysics. Springer, Heidelberg.

Journal of Modern Physics, 2020, 11, 407-431
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Cosmic Expansion: The Dynamic Force Source for All Planetary Tectonic Movements

Jian'an Wang

Department of Physics, Shenzhen University, Shenzhen, China
Email: wja@szu.edu.cn

How to cite this paper: Wang, J.A. (2020) Cosmic Expansion: The Dynamic Force Source for All Planetary Tectonic Movements. *Journal of Modern Physics*, **11**, 407-431.
https://doi.org/10.4236/jmp.2020.113026

Received: February 7, 2020
Accepted: March 14, 2020
Published: March 17, 2020

Abstract

In this paper, the relationship between the geodynamics and cosmic expansion is analyzed and demonstrated from the basic physical principles and various natural phenomena, and the conclusion is drawn that cosmic expansion is the dynamic force source of all planetary tectonic movements in the universe. The energy for the formation and change of landscape and the energy of earthquakes and volcanic eruptions all come from the cosmic expansion. With the cosmic expansion, the energy density of space is decreasing, the atoms and molecules of all matter in the universe are growing, and the magma is expanding and producing gases. As the earth's internal pressure rises, the mechanical energy that accumulates within the magma forms the driving force of the earth's various tectonic movements, and the release of these energy and matter (expansive magma and high-pressure gases) leads to the formation and the changes of the landform (such as orogenesis, epeirogenesis, the formation of the earth's plates, the Earth expansion, the seafloor spreading, and the continental drift), as well as to earthquakes and volcanic eruptions. In this paper, the causes of all kinds of earthquakes, especially deep focus earthquakes, are given and almost all known seismic phenomena are explained, the basic principle and method of earthquake prediction are given, and the direction is pointed out for the elimination of earthquakes and the utilization of earthquake energy. Based on the same principle of physics, this paper also shows that the Ice-Age is caused by the acceleration of the speed of the motion of the solar system relative to the Milky Way in certain regions of the Milky Way. The greater the speed of the solar system relative to the Milky Way, the greater the drop in Earth surface temperature.

Keywords

Geodynamics, Tectonic Movement, Seafloor Spreading, Continental Drift, Orogenesis, Epeirogenesis, Deep Focus Earthquakes, Earthquake Prediction, Ice Age

1. Introduction

There have been numerous theories of geodynamics in the past 100 years, and at present there are mainly the following hypotheses [1]:

1) Hypothesis on the Earth contraction;

2) Hypothesis on the Earth expansion;

3) Hypothesis of the Earth pulsation;

4) Hypothesis on the Earth rotation;

5) Hypothesis on the surge tectonic;

6) A model of layer-block tectonic thermal-upwelling and the Earth's rotation.

The academician of Chinese Academy of Sciences Ma Zhongjin [1] said a reasonable geodynamic hypothesis must satisfy at least three conditions:

1) It can explain the global tectonic characteristics, spatial distribution laws and tectonic evolution process;

2) The dynamic factor it depends on has sufficient energy and its action mode can reasonably explain the characteristics of the tectonic deformation field;

3) It conforms to the basic principles of physics and the physical and chemical properties of earth interior materials.

In terms of the three conditions mentioned above, no hypothesis has been perfect so far.

In the 1960s, geoscience went through a far-reaching revolution. Plate tectonics led us to realize that the surface of the planet on which we are living was divided into rigid plates that have been moving, and that the earth is a dynamic planet, driven by a deep force that causes earthquakes and volcanoes, uplifts mountains and makes the ocean floor to spread, and gradually changes the shape of the earth's surface, thus shaping the spectacular natural landscape on the surface and bringing rich mineral resources to human civilization. However, the revolution is not over, because the source and nature of the driving force behind the planetary tectonic movement are unclear [2].

To sum up, the current geodynamic models are not perfect mainly because the driving force or the energy source for the planetary tectonic movement has not been found.

The purpose of this study is to provide the driving force or energy source for the planetary tectonic movements and to propose a complete geodynamic model. The model not only conforms to the basic principles of physics but also can reasonably explain the various geological tectonic movements, such as orogenesis, epeirogenesis, the formation of the plates, the expansion of the earth, seafloor spreading and continental drift, volcanic eruption and the causes of various earthquakes, especially the deep focus earthquakes, basic principle and method of earthquake prediction. The model can also point out the direction for the future elimination of earthquake and the exploitation and utilization of earthquake energy.

2. The Relationship between Geodynamics and Cosmic Expansion

From the formula (5.31) [3], it can be seen that the orbital radius of the electrons

around the nucleus of the atoms or ions that make up the object decreases with the increase of the velocity of the object in ether. Thus, as the cosmic expansion decelerates, or as the speed of the objects relative to the cosmic etheric system decreases, all the atoms and ions (ions containing orbital electrons) that make up the objects grow larger. So as the cosmic expansion decelerates, the atoms or ions that make up all the earth's material, including the solid crust, underground magma, water, and air, are expanding.

Because the underground magma is in a high-temperature fluid or liquid state, the magma atoms are mostly in an ionic state. As the magma atoms or ions grow larger, the distance between atoms or ions becomes larger and the attraction between atoms or ions becomes smaller, leading to the volatile matter of the magma to release gases. So the underground magma is bubbling and expanding like a foaming foam that continues to produce gas and expand. Therefore, as the cosmic expansion decelerates, the internal hydraulic and pneumatic pressure of the Earth will continue to rise, and the elastic potential energy accumulated in the Earth's magma provides the driving force for the various tectonic movements of the Earth, and the release of the energy and matter (foaming magma and compressed gases) have led to the formation and change of landforms (such as the orogenesis, the epeirogenesis, the formation of plates, the expansion of the earth, the seafloor spreading and the continental drift), and to earthquakes and volcanic eruptions. So the cosmic expansion is the source of the driving force or the source of the energy for the tectonic movements of the earth and all the planets in the universe.

From the formation of natural diamonds and various crystalline minerals in the earth's crust, and from the evidence that meteorites generally have a mineral density greater than earth's, we can learn that all atoms are growing larger as the universe expands. From the fact that there are gases from the volcanic eruption and the fact that the more recent the eruption is the less dense the volcanic rock is, we can learn that the underground magma is continuously producing gases and expanding. From the formation, maintenance and change of earth atmosphere, we can learn that underground magma is continuously producing gases. From a great deal of factual evidence from the geometry, geology, astronomy and geophysics of continental forms, we can learn that the earth has been expanding. This geodynamics model gives reasonable explanations of various tectonic movements, such as orogeny, epeirogenic movement, seafloor spreading, continental drift, volcanic eruptions, and reasonable explanations of the causes of various earthquakes (including deep-focus earthquake) and various seismic phenomena.

2.1. From the Formation of Natural Diamonds and Various Crystalline Minerals in the Earth's Crust, and from the Evidence That Meteorites Generally Have a Mineral Density Greater than Earth's, We Can Learn That All Atoms Are Growing Larger as the Universe Expands

Why can't the lava from the volcano now form crystalline minerals such as diamond, gem, granite and so on after solidification, but light volcanic rocks? Ac-

cording to formula (5.31) [3], all atoms will increase with the decelerating expansion of the universe, so the atoms and molecules of magma are much larger than those of hundreds of millions or billions of years ago, and under the normal pressure of one atmosphere, minerals such as diamond, gem and granite cannot be formed by natural solidification. Crustal minerals such as diamonds, gemstones and granites were formed by magma cooling and solidifying at different ages during the earth's crustal formation process. Diamonds were formed billions of years ago by natural cooling and solidification of magma on the earth's surface or at shallow crust at about one atmospheric pressure. This is why we can find or mine diamonds on the surface of the earth or in the shallow layers of the earth's crust. Natural diamonds must be found in ancient continental plates such as Africa and Siberia, not in young craters or on young continental or oceanic crusts.

In addition, meteorites have a high probability of containing diamonds because of their small size and quick solidification (the time of solidification is very early).

So the fact that natural diamonds and various crystalline minerals formed on the earth's surface, and that meteorites have a higher mineral density than earth, provides evidence that atoms keep getting bigger as the universe expands.

2.2. From the Fact That There Are Gases from the Volcanic Eruption and the Fact That the More Recent the Eruption Is the Less Dense the Volcanic Rock Is, We Can Learn That the Underground Magma Is Continuously Producing Gases and Expanding

The fact that underground magma continues to produce gases from volcanic eruptions can be verified by the fact that carbon dioxide, hydrogen sulfide, sulfur dioxide, hydrogen chloride, hydrogen fluoride, nitrogen, argon, methane, carbon monoxide and other gases are accompanied by the eruptions.

We can measure the specific gravity of volcanic rock in the vicinity of the crater at different ages, indicating that the closer the eruption is, the less dense the rock is. As shown in **Figure 1**, the density of 6-edged basalt volcanic rocks in the ancient crater of the Changle Northern Rock in Shandong province of China 18 million years ago is significantly larger than that of the Shishan volcanic group in Haikou city, Hainan province of China about 13,000 years ago, as shown in **Figure 2**.

As the universe expands, the atoms and ions of the underground magma keep getting bigger, the magma keeps producing gases at an accelerating rate, so the magma has more and more bubbles and the bubbles are getting bigger and bigger, and the magma keeps expanding. That's why the closer the age of the eruption, the less dense the rock.

2.3. From the Formation, Maintenance and Change of Earth Atmosphere, We Can Learn That Underground Magma Has Been Producing Gases Continuously

As the universe expands, the atoms and molecules of everything in the universe

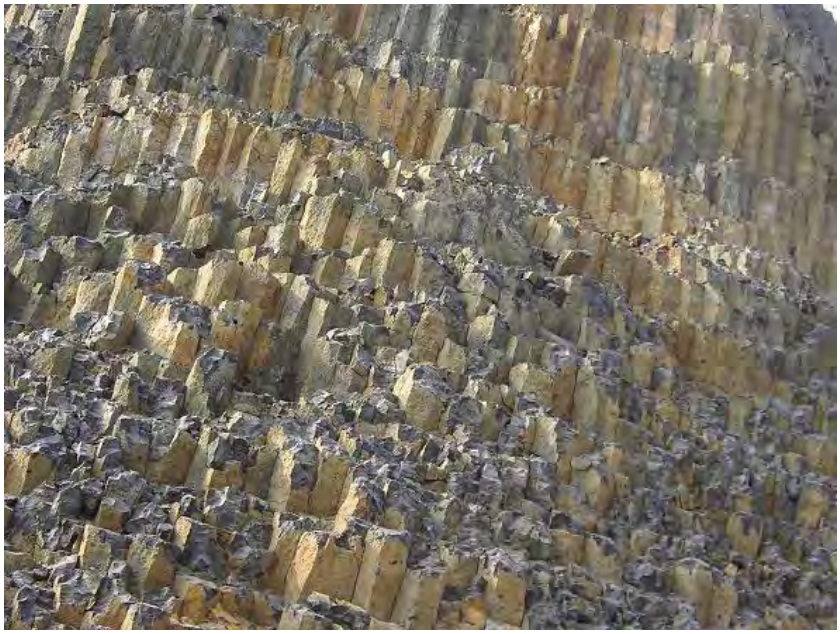


Figure 1. 6-angled basalt volcanic rocks from the ancient crater 18 million years ago in Changle North Rock, Shandong Province.

Figure 2. Volcanic rocks from Shishan volcanic group about 13,000 years ago, in Haikou, Hainan province.

keep getting bigger. In the early days of the earth, small molecules of some volatile materials in the magma, such as hydrogen and helium, broke free from the magma and became free molecules. These free molecules gathered on the surface of the planet after leaving the magma to form the planet's primeval atmosphere by the gravity. Since atmospheric molecules obey Maxwell's law of speed distribution, there is always a certain percentage of gas molecules at every moment that exceed the second cosmic velocity, so the atmosphere is constantly losing molecules to space. The earth's atmosphere would not have lasted for billions of years without the constant replenishment of gases. It is because the gases re-

leased by the underground magma continue to replenish the air lost in the atmosphere that the Earth's atmosphere has been able to remain so long. As the cosmic expansion continues to slow down, the magma atoms continue to grow, the weight and size of the molecule released from the magma keep increasing and the composition of the atmosphere is constantly changing. In the early days of the Earth, underground magma mainly released small molecule gases such as hydrogen and helium, and now it mainly releases large molecule gases such as nitrogen, carbon dioxide, sulfur dioxide and hydrogen sulfide. A paper by the university of California, published in Science on April 6, 2018 [4] claimed to have found that up to 26% of nitrogen in the ecosystem came from the earth's lithosphere bedrock (this paper suggests from underground magma).

The disappearance of the atmosphere of the moon and mars is due to the fact that the mass of the moon and Mars is much smaller than that of the Earth (the masses of the moon and Mars are about 1/81 and 1/10 of that of Earth, respectively), and that the underground magma has completely solidified and stopped producing gas.

2.4. The Evidence of Earth's Expansion

In the article "the theory of earth expansion, its development and its main facts" [5], a large number of facts were provided for the continuous expansion of the earth from the geometry of the continental form, geology, astronomy, geophysics and other aspects.

2.5. Explanations of the Formation and Variation of the Earth landforms by the Geodynamics Model Based on Cosmic Expansion

The formation of earth landforms such as the land, the sea, the mountain chain, the cave, the mountain and the basin are caused by cosmic expansion. The formation of the earth's plates, the seafloor spreading, and the continental drift are also caused by cosmic expansion.

As the Earth's surface continues to cool, the Earth's surface slowly formed a solid monolithic shell, the lithosphere. The outer layer of the lithosphere is solid, and the further inside, the higher the temperature, the softer. The continuous gas production of underground magma causes the expansion rate of underground magma to be greater than that of solid lithosphere. Before the formation of the Earth's plates, or before the broke up of the Earth's lithosphere, the Earth was like a closed, thin-walled container filled with hot water at 100˚C and continuously heated. With the expansion of underground magma and the accumulation of high-pressure gases under the lithosphere, these foaming magma and high-pressure gases constantly changed the shape of the lithosphere both leading to violent orogenic and epeirogenetic movements.

With the cosmic expansion the gases produced within the Earth grew, and these gases gathered beneath the lithosphere formed a gas layer. Because of the inhomogeneity of the lithospheric material structure, the rate at which each part

cools and solidifies will vary, and those relatively weak local areas will be uplifted by the gas pressure to form mountains, while those with no or relatively few uplifts will form valleys, basins, or plains. The gases that partially jacked up the lithosphere to form the mountains would eventually escape the earth and leave an exhaust passage (the original cave) underground. Thus, one end of the cave must be connected to the atmosphere and the other to the mantle.

With the continuous cooling of the earth's surface, the lithosphere has been constantly thickening. When the lithosphere reached a certain thickness, the pressure required to rupture the lithosphere is less than that required for the partial uplift of the lithosphere, the foaming magma and high-pressure gas finally ruptured the lithosphere into several large plates.

Because the sudden rupture of the lithosphere caused the sudden expansion of the earth and the sudden separation of the primitive plates formed after the rupture of the lithosphere, the surface area of the earth suddenly increased a lot. Because the curvature of the primitive plates formed after the rupture of the lithosphere is smaller than that of the expanded earth's surface, a layer of gas accumulated under the primitive plates after the rupture of the lithosphere. Because there is a gas layer under the primitive plates, the primitive plates have been floating on the magma and form the continents.

Because these new crusts, formed by magma cooling from the gap between the primitive plates, are relatively low in topography, they constitute the ocean floors. Because these newly formed crusts (ocean floors) are formed by magma cooling exposed by the gap between the primitive plates formed by lithosphere rupture, there is a great difference in geological age between continental crusts and oceanic crusts.

Because there are fragments in the middle of the gap between the primitive plates during the rupture of the lithosphere, these fragments have been remaining in their original position during the seafloor spreading, thus becoming island or continental fragments.

Because the sudden rupture of the lithosphere lead to a massive release of underground gas, including water vapor, air and water vapor in the atmosphere suddenly increased. The cooling of the water vapor formed the original ocean.

Because the oceanic crusts are thin, the tectonic movement in the vertical direction of the earth after the rupture of the lithosphere has been being concentrated on the oceanic crusts mainly.

As the universe continues to expand, the underground magma continues to produce gases and causes the earth to continue to expand, causing the newly formed ocean crusts to continue to burst, and the magma continues to gush out at the fractures of the ocean crusts to form and renew mid-oceanic ridges in the ocean. As the underground magma continues to expand, the ocean floor expands on both sides of the mid-oceanic ridge, and underground magma gushes out to form new ocean floor. This is how the ocean floors spread and continents drift. Thus, the closer to the mid-oceanic ridges, the younger the ocean floor.

It is the expanding force of the underground magma, not the convection of

the earth's mantle, that cause the seafloor to spread, the mantle plumes and plate subductions do not exist. Because the earth is expanding, plates are moving away from each other, so plate squeezes and collisions do not exist. This shows that the earthquake is not caused by plate squeeze or collision.

Because the gaps between plates are cemented with solidified magma, the plates tear away from each other as the earth expands, starting at a point in the gap. So the path of the continental drift is an arc.

In the early days of the ocean floor spreading, the ocean crust (the gap between the primitive plates) was narrow and thin, although the water above the ocean crust has weight, but water has no stress to resist deformations, so as the magma pressure inside the Earth increased, some ocean ridges and the surrounding ocean crust would be uplifted by underground magma to form new land and mountains above the sea level. This is why fossils of ancient marine animals and plants have been found in some mountains, such as the Himalayas.

Because under the force of underground magma expansion the earth continues to expand and the plate curvature becomes smaller and smaller, the gap between the two plates will continue to bulge outward along the radius of the earth, causing these mountain ranges formed along the gap between two plates, such as the Himalayas, to rise.

As the earth expands with the seafloor spreading, the curvature of the earth's surface becomes smaller and smaller. Therefore, the older the plate, the greater the curvature. Since the continental crust is older than the oceanic crust, the boundary between the continental crust and the oceanic crust forms the trench.

The continuous gas production of underground magma makes the expansion rate of underground magma larger than that of the water on earth, so the ratio of the volume of water, including ice, to the volume of the earth has been decreasing and the ratio of land area to ocean area has been increasing.

2.6. Explanation of Volcanic Eruptions by the Cosmic Expansion Geodynamic Model

As the universe continues to expand, the magma atoms inside the Earth continue to grow and the magma continue to produce gases. Bubbles gathered by these gases are constantly floating upward under the buoyancy of the magma and converge in the plate fault zones and craters. In an active volcano, magma containing bubbles is constantly spewing out of the crater. In an extinct volcano, when the pressure of the bubbling magma is greater than the threshold, the magma breaks through the crater blockage and erupts.

2.7. Explanations of the Causes of Various Earthquakes by the Cosmic Expansion Geodynamic Model

According to this geodynamic model, earthquakes are caused by cosmic expansion that causes the underground magma to continuously separate out gases.

As the universe continues to expand, the magma atoms continue to grow, leading to the continued generation of gases in the magma (primitive air, now

mainly nitrogen and carbon dioxide). Since these gases are produced in the mantle and core, and at high temperatures in the mantle and core gas molecules are ionized. These bubbles, driven upward by the buoyancy of the magma, accumulate beneath the plates and in cracks in the fracture zones. These high temperature and high pressure charged gases, which gather under the lithosphere, are constantly looking for a breakthrough to release. Because the solidification layer at the junction between two plates is relatively weak, it is easy to be broken, resulting in the sudden release of this high temperature and high pressure charged gases (explosion)—earthquake.

On June 12, 2008, after Chinese Wenchuan earthquake, "People's Weekly", titled "Wang Shencun: A lonely fan of scientific research of the field of earthquakes", reported on his thirty-two years of silent research and exploration of the mysteries of earthquakes, and published the full text of chapter 1 - 4 of his book "reveal the secret of nature-secrets of earthquakes". "Reveal the secret of nature-secrets of earthquakes", also known as "new theory of earthquakes" or "earthquake gas explosion theory", was proposed by author Wang Shencun shortly after the 7.8-magnitude Chinese Tangshan earthquake on July 28, 1976, according to the phenomenon of gas explosion occurring during the earthquake: light (flash), sound (explosion), gas (gas breaking through the ground and rising into the air), and quake (earthquake), and after many years of difficult research and exploration and simulation tests, a more systematic theory of seismic gas explosion was preliminarily completed in October 1981, which opened up a new way to solve the world problem of earthquake [6].

The cause of earthquakes derived from this Cosmic Expansion Geodynamic Model is in perfect agreement with Mr. Wang Shencun's Earthquake Gas Explosion Theory. Since this Cosmic Expansion Geodynamic Model is derived from the conclusion of the Modified Special Relativity [3], and Mr. Wang Shencun's Earthquake Gas Explosion Theory is summed up by a large number of seismic phenomena, the credibility of the conclusion that earthquake is caused by the explosion of high temperature and high pressure charged gas is very high.

2.7.1. The Mechanism of Natural Earthquakes

As the universe expands and the underground magma continues to produce gases, the pressure on the high-temperature and high-pressure charged gases that accumulate beneath the lithosphere increases. When the pressure is greater than the strength of the surrounding rock mass, a gas explosion occurs. This is how natural earthquakes occur.

2.7.2. The Mechanism of Artificial Earthquakes

The explosion of the high temperature and high pressure charged gases stored under the lithosphere can also be triggered by man-made factors, such as storage or release of water from reservoirs, underground nuclear experiments or large equivalent TNT underground explosions, which is the mechanism of artificial earthquakes.

2.7.3. The Mechanism of Deep-Focus Earthquake

The deep-focus earthquake is the earthquake with a depth of more than 300 km, the deepest known deep-focus earthquake with a depth of 720 km. Studies have shown that rocks below tens of kilometers are in soft plastic state. Below 720 km the temperatures is above 1100˚C degrees and the rock is basically fluid. As the universe expands, magma atoms expand and separate gas molecules (actually ions) that gather into bubbles. These bubbles move upwards under the buoyancy of the magma. Because the farther away from the center of the earth, the lower the hydraulic pressure of the magma, the lower the temperature, the thicker the magma. Therefore, in the process of the bubble floating upward, the resistance encountered is increasing, and the surface tension of the bubble is also increasing, and the hydraulic pressure outside the bubble is decreasing. Bubbles get bigger and bigger as they rise upward (more and more bubbles merge). Because the earth's radius is very large and the velocity of the bubble floating up is very small (because the magma is very dense), the process of the bubble floating up is very long. In the long process of bubble floating, the elastic potential energy gathering in the bubbles will be larger and larger, and when the pressure inside the bubble is greater than the sum of the hydraulic pressure of the external magma and the surface tension of the bubble, the bubble will explode. It can be seen that the deep-focus earthquake is the explosion of high pressure bubbles in the underground magma.

2.8. Explanations of Seismic Phenomena by the Cosmic Expansion Geodynamic Model

In the long-term practice, the ancient Chinese people realized that earthquakes have precursors, and left a wealth of records about earthquake precursors. Liang Guanghe's article "new knowledge of earthquakes" [7] and Yue Zhongqi's blog "the causes that each airs his own views for strong seismicity and the prediction" [8] list a large number of natural phenomena that have been observed before, during and after earthquakes based on seismic investigations and research. The author summarized the above seismic phenomena and explained them one by one by using this cosmic expansion geodynamic model:

2.8.1. Explanations of the Main Natural Anomalies before Earthquakes, the Accompanying Phenomena during Earthquakes and the Phenomena after Earthquakes

Because there will be a lot of high temperature and high pressure charged gas gathering and activity in the underground of the earthquake area before the earthquake occurs, there will be many abnormal natural phenomena. The accumulation and slow release of high temperature and high pressure charged gases will not only affect the nearby electromagnetic field but also change the temperature and weather. In addition, the accumulation and release of these high temperature and high pressure charged gases will also lead to local underground pressure changes, resulting in the formation deformation and abnormal sound. The variation of the underground pressure and the deformation of the formation

will lead to the change of the groundwater level.

1) Groundwater anomalies: Groundwater includes well water, spring water, etc. The main anomalies are muddy, bubbling, churning, temperature rise, discoloration, flavor change, sudden rise, sudden drop, well hole deformation, spring source suddenly dried up or gushing, etc.

Explanation: The anomaly of groundwater is caused by the accumulation and release of underground high temperature and high pressure charged gases in the preseismic area. The slow release of high temperature and high pressure charged gases accumulated underground can cause particles of underground minerals or soil dust to enter the water and cause well or spring water to become muddy, bubbling, churning, warming, discoloring, or flavoring. Because the change of the high temperature and high pressure charged gas pressure in the underground will change the hydraulic pressure of the groundwater and cause the deformation and movement of the formation, it will lead to the sudden rise or sudden drop of the well water, the deformation of the well hole, sudden depletion or emission of the spring source, etc.

2) Biological anomalies: Before the earthquake, in addition to animals have abnormal behavior, some plants also have abnormal response, such as inappropriate season of germination, flowering, fruiting or a large area of wither and unusually exuberant.

Explanation: The abnormality of animals is mainly due to the underground high temperature and high pressure charged gas release in the pre-earthquake zone, which leads to the interference of the geomagnetic field, the abnormal sound and the peculiar smell. The plant anomaly is mainly due to the slow release of high temperature and high pressure charged gases in the in the pre-earthquake zone. Because the temperature of the gases is very high, the release of these gases can change the temperature of the soil and air in the zone, resulting in the germination, flowering and fruiting of plants during the inappropriate season. Because the main components of seismic gases are nitrogen and carbon dioxide, and these elements are the nutrients that plants need, the slow release of these gases can cause plants to become unusually lush. But if the released gas is too thick, it will cause the plant to wither in large area.

3) Weather anomalies: Before earthquakes, the weather is often abnormal. The main abnormal phenomena of the weather are: sultry, long drought or steady rain, yellow fog everywhere, dark daylight, strange wind raging, hail in June and so on.

Explanation: The gathering and slow release of underground high temperature and high pressure charged gases in the earthquake zone before the earthquake will certainly have an impact on the weather in the earthquake zone. The sultry weather is mainly caused by the release of the underground hot gases and the evaporation of the groundwater caused by the release of the underground hot gases. If there isn't much ground water in the zone, then the released hot gases will lead to a prolonged drought, and if the zone is rich in groundwater, the released hot gases will lead to the evaporation of groundwater, resulting in

steady rain. Since the charged gases can absorb dust particles during the underground flow, the release of these charged gases into the upper air may cause the high-altitude water vapor to form the hailstorm, so there may be hailstorms in June. The release of these charged gases containing dust particles from the ground can also lead to yellow fog and dark sunlight. The strange wind raging phenomenon should be caused by the rising hot air from the release of the high temperature and high pressure charged gases stored underground.

4) The emergence of the Earthquake Cloud: Earthquake Cloud is a kind of cloud in non-meteorological classification that indicates earthquakes. At present, the academic community is skeptical about the Earthquake Clouds, believing Earthquake Clouds to be pseudoscience, because, according to the existing seismic theory, earthquakes are caused by the accumulation and release of crustal plate stress due to crustal movement, and the slow movement of the crusts is not related to the meteorological phenomenon for a short time. There are only folk enthusiasts to explore the earthquake cloud in China, Japan, Indonesia and other countries.

Explanation: Because a large amount of high-temperature and high-pressure charged gases will accumulate underground in the earthquake area before the earthquake occurs, when these high-temperature and high-pressure charged gases accumulate to a certain amount, they will overflow from the ground (because when the pressure exceeds a certain value the surrounding rock mass that gathers these high-pressure gases will leak). Because the charged gases can absorb dust particles during underground flow, the release of a large amount high-temperature charged gases from the ground above the surrounding rock mass (possibly source) where these gases are gathered into the atmosphere can lead to the generation of this particular cloud-earthquake cloud. As a result, many earthquake clouds are strip-shaped and point to the epicenter at one end. If these gases are released from multiple outlets (potential earthquake sources) in the fault zone, an earthquake cloud as shown in **Figure 3** will result. The emergence of the earthquake cloud is similar to the situation that the river level has exceeded the warning line and began to overflow (the possibility of a dam broke is very high). Therefore, the emergence of the earthquake clouds indicates that the probability of a large earthquake is very high.

5) Underground sound anomaly: The underground sound anomaly refers to the sound from underground before earthquakes that is like the thunder of cannon, or the sound of heavy vehicle driving, or the sound of the strong wind surge and so on.

Explanation: Earthquakes are caused by the explosions of high temperature and high pressure charged gases in the ground. Because underground sound travels faster than the shock wave from gas explosions, the sound of cannon thunder will be heard before the earthquake. A large amount of gases released by a gas explosion flowing underground, especially in underground caves, can produce the sound of heavy vehicle driving, or the sound of the strong wind surge.

Figure 3. Earthquake clouds.

6) Abnormal ground light: Abnormal ground light refers to the light from the ground before an earthquake, its color is varied, rare mixed colors in daily life can be seen, such as silver blue, white purple, but mainly red and white; its shape is also varied, there are ribbon, ball, columnar, diffuse, etc. Generally, ground light appears in a wide range, mostly in the first few hours to a few minutes before the earthquake, lasting several seconds.

Explanation: A large amount of high temperature and high pressure charged gases will be gathered underground in the earthquake area before the earthquake occurs. The ground-light phenomenon should be the discharge on the ground of these charged gases when they flow from the ground to the air before the earthquake.

7) Abnormal ground gas: Abnormal ground gas refers to the fog from the ground before the earthquake. This fog, with a variety of colors, such as white, black, yellow, and sometimes colorless, often occurs within a few days to a few minutes before the earthquake, often accompanied by a strange smell, sometimes accompanied by sound or high temperature.

Explanation: The ground gas phenomenon is caused by the release of high-temperature and high-pressure charged gases gathered underground before the earthquake. Because the charged gas can absorb the dust particles of various substances it passing through during the underground flow, the odour should be caused by the dust particles of certain substances. Since the gases released are very hot, of course there will be high temperature phenomenon accompanied by. The rapid flow of a large amount of gas underground certainly will make some sounds.

8) Abnormal ground movement: Abnormal ground movement refers to the shaking of the ground before the earthquake. Before the earthquake, people

sometimes feel the ground shaking, which is different from the earthquake. The shake is so slow that seismometers often miss it, but many people can feel it.

Explanation: Abnormal ground movement is a small shaking caused by the high pressure gas accumulated underground in the seismic area before the earthquake, which pushes up the local strata and causes the strata to be unstable.

9) Abnormal ground bulge: Abnormal ground bulge refers to bulges formed on the ground before the earthquake

Explanation: The abnormal ground bulge is caused by the high pressure gases accumulated underground in the earthquake area before the earthquake.

10) Electromagnetic phenomena associated with earthquakes: To sum up, the electromagnetic phenomena associated with earthquakes are the production of strong current and the sudden change of electromagnetic field near the quake zone. The electrical current generated by an earthquake is very strong. A magnitude 6 earthquake can generate a current of up to 100,000 amperes and a magnitude 7 earthquake can generate a current of up to 1 million amperes [7]. The maximum current on a typical household meter is usually no more than 60 amperes. How does such a strong current come about?

Explanation: Because the earthquake is formed by the explosion of high temperature and high pressure charged gases gathered underground, there will be a rapid underground flow of a large number of ionic gas molecules when the earthquake occurs, thus forming a strong current and producing a strong magnetic field. The strong magnetic field generated by an earthquake can cause a sudden change in the electromagnetic field near the earthquake area.

11) Large earthquakes are often the metallogenic process of underground metalliferous minerals. Scientific research showed that earthquakes can precipitate gold deposits.

Explanation: Because seismic gases are produced in the mantle and core, the temperature of these gases is very high. Large amounts of hot seismic gases flowing through the earth's crust during a great earthquake can cause passing gold particles to melt and solidify into gold nuggets, or cause metalliferous minerals to melt and produce reduction reactions to form metals.

12) During an earthquake, people first hear the sound and then feel the quake.

Explanation: The earthquake is caused by the explosion of high-temperature and high-pressure gases formed by the accumulation of charged gas molecules separated by underground magma. Because the sound wave travels faster in the rock than the shock wave produced by the explosion of high-pressure gas, people first hear the sound and then feel the quake.

13) A great earthquake may immediately darken the sky

Explanation: The earthquake was caused by the sudden release (explosion) of high-temperature and high-pressure gases formed by the accumulation of charged gas molecules separated by underground magma. The charged gas traps dust particles as it travels underground. A great earthquake can immediately darken the sky because of the sudden release of a large amount of gas adsorbed

with dust particles from the ground into the atmosphere.

14) The air temperature may begin to drop immediately during or after an earthquake

Explanation: Because the earthquake is caused by the sudden release (explosion) of high pressure gas underground, and the sudden decompression release of high pressure gas will absorb the heat of the surrounding material, the temperature in the earthquake area will suddenly drop during the earthquake.

15) Heavy rain or snow may fall in hours after a strong earthquake

Explanation: Because the earthquake is caused by the sudden release (explosion) of high pressure gas underground, and the sudden decompression release of high pressure gas will absorb the heat of the surrounding material, the temperature in the earthquake area will suddenly drop during the earthquake. A sudden drop in temperature can cause water vapor in the air to quickly condense into water droplets, or snowflakes. So it may rain or snow in a few hours after a great earthquake.

16) There is underground wind in an earthquake

Explanation: Because the earthquake is caused by the sudden release (explosion) of high pressure gas underground, the flow of the gas from the earthquake forms the underground wind.

17) Before and after an earthquake, there may be natural gas ejection, explosion, and combustion

Explanation: Because the earthquake is caused by the sudden release (explosion) of high-temperature and high-pressure gases formed by the accumulation of charged gas molecules separated by underground magma, if the source is below the natural gas field, the earthquake can cause a large amount of gas to be released into the natural gas field, so the gas reserves and pressure in the natural gas field will increase. Because a large amount of charged gas enters into the natural gas field when an earthquake occurs, if this gas mixture in the natural gas field is ejected into the atmosphere, it could lead to spontaneous combustion or explosion.

18) During an earthquake, a fire, even a huge fire, may break out in a city

Explanation: Because earthquakes are caused by the sudden release (explosion) of high-temperature and high-pressure charged gases formed by the accumulation of charged gas molecules separated by underground magma, the encounter of these high-temperature charged gases in the atmosphere with leaking liquefied petroleum gas, gasoline or other flammable gases or liquids can lead to the combustion or explosion of these flammable gases and liquids. Therefore, there may be a fire or even a huge fire in the city during the earthquake.

19) Earthquakes can trigger forest fires (the author speculates)

Explanation: There are frequent forest fires on Earth [9], such as the Amazon forest fire in August 2019, the forest fire in California in September 2019, and the forest fire in Victoria, Australia, in November 2019. How did these forest

fires come about? Noting that California and Amazonian forests are located in the Pacific seismic belt, Australia's Victoria is also the region where the earthquake occurred. Because earthquakes are caused by the sudden release (explosion) of high-temperature and high-pressure charged gases formed by the accumulation of charged gas molecules separated by underground magma, if the earthquake occurs in the forest area, especially in the forest area where there is no rain for a long time, the instantaneous release of a large amount of high temperature charged gas is completely likely to ignite the hay, fallen leaves and dead trees of the forest from many large areas, thus leading to the occurrence of forest fire. It can be seen that the forest fire caused by the earthquake has the characteristics of igniting in a large area and many places at the same time, and starting very quickly, it has no time to extinguish.

20) In strong earthquakes, local mountain soil and rock are thrown out by gas, mountain collapse and earth crack, and the coseismic surface soil is broken (tensile and shear type)

Explanation: Because earthquakes are caused by the explosion of high temperature and high-pressure charged gas, in strong earthquakes, local mountain soil and rock will be thrown out by the gas, mountain collapse and earth crack. Because the shock wave produced by the underground gas explosion is spherical longitudinal wave with the explosion point as the center, the rupture of the coseismic surface soil is tensional and shear type.

21) There may be a huge increase in the underground natural gas reserve after the earthquake

Explanation: If the earthquake source is below the natural gas field, because the earthquake may cause a large amount of gas to be released into the natural gas field, the gas reserve in the natural gas field may has a huge increase, but the total content of natural gas will not increase, and the nitrogen and carbon dioxide content in the natural gas will increase a lot.

22) Earth's rotation rate may accelerate after a great earthquake

Explanation: Since a large quantity of high-pressure gas that accumulates beneath the earth's crusts are released into the atmosphere during a great earthquake, the crust sinks slightly and the earth's moment of inertia becomes smaller. Because of the conservation of angular momentum on the earth's rotation, the angular speed of the earth's rotation will accelerate slightly.

23) After a great earthquake, the ground has a large and wide range of horizontal displacement and settlement

Explanation: Because earthquakes are caused by the sudden release (explosion) of high-temperature and high-pressure gases formed by the accumulation of charged gas molecules separated by underground magma. Before the earthquake, the local crust of the earthquake area is lifted up by the high pressure gas, and after the earthquake the high pressure gas is released from the underground into the atmosphere, so after a great earthquake, the ground in the earthquake area will settle down greatly and widely. Due to the geological structure of the

ground in the process of subsidence will also produce horizontal movement.

2.8.2. Explanations to the Laws of Earthquake Occurrence

1) The sites of the earthquake are located in the weak zones and deep fault zones of the crustal geological material

Explanation: Earthquakes are caused by the sudden release (explosion) of high temperature and high pressure gases formed by the accumulation of charged gas molecules separated by underground magma. Because the weak zone of geological materials can be easily destroyed by the high pressure gas, the sites of the earthquake are located in the weak zones of crustal geological material. Because the deep fault zones tend to gather bubbles of gas separated by underground magma (the bubbles float upward by the magmatic buoyancy, they can easily enter the cracks in the deep fault zone during the process of floating), and the fault zones are prone to rupture, the sites of the earthquake are located in the deep fault zones.

2) Where many weak or small earthquakes occur, no strong or great earthquake will occur

Explanation: Because in the place where many weak earthquakes and small earthquakes occur, the accumulated underground high pressure gas is released in time, and the energy gathered by the underground high pressure gas will not be large, so there will be no strong or great earthquake.

3) The focal depth of medium and small earthquakes varies greatly, while the focal depth of strong and great earthquakes is usually shallow (within 25 km depth)

Explanation: Because the deeper the underground, the higher the temperature, the softer the crust, the smaller the energy accumulated before the explosion of high-pressure gas; in addition, the deeper the underground pressure (mainly the pressure formed by the formation weight), the smaller the pressure difference between the bubbles formed by high-pressure gas and the outer space; therefore, the deeper the underground, the more difficult it is to form a strong or great earthquake. Within 25 km of depth, the underground temperature is not high, and the strength of underground rock is generally relatively large. In addition, the underground pressure within 25 km is relatively small. Therefore, if the surrounding rock mass with the accumulation of underground gas is relatively strong, the pressure difference of the high pressure gas gathered in the rock bed before explosion can be very large relative to the outer space, and the elastic potential energy accumulated can be very high. Therefore, the source depth of strong earthquakes and great earthquakes is usually shallow (within 25 km depth). If the surrounding rock mass for the accumulated underground gas is not strong enough, then within 25 km depth can also produce small and medium earthquakes.

4) The sequence of earthquake types is diverse, such as foreshock main shock aftershock type, main shock aftershock type, double main shock type, etc

Explanation: Earthquakes are caused by the sudden release (explosion) of high

temperature and high pressure gas formed by the accumulation of charged gas molecules separated by underground magma. Foreshocks, main shocks and aftershocks are all formed by the explosion of this high-pressure gas. Due to the diversity of geological structures in the geological fault zone, that is to say, the surrounding rock masses containing such high temperature and high pressure charged gas are different in size and structure, so the sequence of series explosions generated by the high-pressure gas gathered in the surrounding rock masses is of course diverse.

5) Earthquakes can be repeated in one place, and the repeated periods vary widely among geological regions, for decades, hundreds or thousands of years

Explanation: Because earthquakes occur in fault zones where underground gas is easily accumulated and the fault zone does not disappear as a result of earthquakes, earthquakes can be repeated in one place. After the earthquake, because of the formation of many underground gas channels in the earthquake area, it is temporarily difficult to gather gas below the earthquake area. Over time, the underground magma solidifies in these gas channels, blocking them, and then allows the underground gas to accumulate below the earthquake area. Due to the different time cycles for the blockage of underground gas channels in different seismic areas and the different time cycles of gas reaccumulation, the seismic repetition cycles in different geological areas vary greatly, which can be tens, hundreds or thousands of years.

6) The earthquake area may have seismicity gaps

Explanation: If there is a solid rock bed with a large mass or a solid rock bed connected to the plate as a whole above or obliquely above the earthquake source, the area above the rock bed may form a seismicity gap. Because earthquakes are caused by the explosion of underground high-pressure gas, the bulk of a solid rock bed with a large mass or a solid rock bed connected to a plate as a whole is largely unaffected by shock waves from gas explosions.

7) Earthquakes can cause tsunamis

Explanation: Earthquakes are caused by the explosion of underground high-pressure gas. Because the shock waves from underground gas explosions can be transmitted from the stratum to the ocean, they can cause tsunamis.

8) Coseismic surface fractures observed on sites generally occur in weak strata such as surface soil, weathered soil and coal seams, while coseismic fractures of hard surface rock mass are rarely seen

Explanation: Because the fault of the co-seismic surface is caused by the shock wave generated by the underground gas explosion, and the strength of the shock wave reaching the surface is not enough to break the hard rock mass, but can break the weak strata such as the surface soil, weathered soil and coal seam.

9) The duration of a single earthquake is very short, from seconds to tens of seconds

Explanation: A earthquake is usually caused by a series of explosions of many high-pressure gas pockets that form in a geologic fault zone. Because a single

earthquake is caused by the explosion of a single high-pressure gas pocket, the duration is very short, from seconds to tens of seconds.

10) The attenuation rate of earthquake intensity along horizontal distance is much lower than that along depth

Explanation: Because the earthquake is caused by the explosion of underground high-pressure gas, on the surface of the ground the seismic wave is circle propagation centered on the epicenter, while in the earth it is spherical propagation centered on the seismic source. Since the circumference of the circle is proportional to the radius of the circle and the surface area of the sphere is proportional to the square of the radius of the sphere, the attenuation rate of the seismic intensity along the horizontal distance is much smaller than that along the depth.

11) The macro-epicentre of an earthquake is located in a weak zone of geological materials (such as in or near a basin)

Explanation: Because the earthquake is caused by the explosion of underground high pressure gas, and the weak zone of geological materials has little ability to withstand shock waves, it is easy to be destroyed by shock waves, so the macro epicenter of an earthquake is located in the weak zone of geological material.

12) The vertical acceleration of the epicentre is greater, and the horizontal acceleration at a distance from the epicentre is greater

Explanation: Because the seismic wave is the longitudinal wave propagating along the radius of a sphere with the gas explosion point as the center of the sphere, the epicentre has a greater vertical acceleration and at a distance from the epicentre has a greater horizontal acceleration.

13) The period of calm in each fault zone prior to a strong earthquake is decades, hundreds, or thousands of years

Explanation: Because fractures in the fault zone are vertical, they are particularly good for the accumulation of gas bubbles that float vertically up from the mantle. The larger the space of the gas chamber formed by the fault zone, the greater the strength of the fault zone, the longer the time of collecting the underground gas in the fault zone, the greater the elastic potential energy of the gas gathered, and the stronger the earthquake produced. So each fault zone has decades, hundreds, or thousands of years of calm before a strong earthquake.

14) Casualties in meizoseismal areas are generally caused by mechanical kinetic energy such as building collapse and landslides. Some people feel uncomfortable due to breathing gases and dust during the earthquake, and no one seems to be harmed by the gases

Explanation: Because the gases released by the earthquake are the original gases that makes up the air (air that is not absorbed and converted by animals and plants), the main components are nitrogen and carbon dioxide, which are non-toxic and harmless.

15) The distribution of large seismic zones, large natural gas fields and large fault zones on the earth is consistent

Explanation: Because the large fault zone is easy to gather the bubbles formed by the high temperature and high pressure charged gases separated by the underground magma (because the bubbles will float up under the magmatic buoyancy, and they are easy to enter the large fault zone during the process of floating), and the fault zone is easy to break, so the large fault zones are large seismic zones. And because large amounts of high temperature and high pressure charged gases often accumulate underground in large seismic zones, the gases trapped underground are released slowly for a long time before an earthquake. The release of high temperature and high pressure gases will increase the temperature of the soil and air in the seismic zone, which is good for the growth of plants. Moreover, since the gases separated by the underground magma are mainly ionized nitrogen and ionized carbon dioxide, which are the nutrients needed by plants, and ionized gas is easy to be absorbed by plants, the slow release of these gases will lead to abnormal flourishing of plants. Because earthquakes often occur in the seismic zones, the plants that grow in the seismic zones are often buried underground to form coal fields and natural gas fields. Therefore, the distribution of large seismic zones, large natural gas fields and large fracture zones on the earth is consistent

16) A huge earthquake may cause short-term increase or change in atmospheric pressure

Explanation: Earthquakes are caused by the sudden release (explosion) of high-temperature and high pressure gases formed by the accumulation of charged gas molecules separated by underground magma. A huge earthquake can cause a sudden release of large quantities of high pressure gas from the ground into the atmosphere, resulting in short-term increases or fluctuations in atmospheric pressure.

17) There are a large number of active erupting volcanoes near many fault seismic zones

Explanation: Because fractures in fault zones tend to form mantle passageways to the ground surface, and fault zones are also conducive to collecting and storing gases produced by underground magma, there are a large number of active erupting volcanoes near many fault seismic zones.

3. Discussions

3.1. About Earthquake Prediction

Because there is a high temperature and high pressure charged gas accumulation process underground in the earthquake area. The accumulation of these high-temperature and high-pressure charged gases will inevitably cause the underground pressure to rise and produce gas leakage. The rise of the underground pressure will cause the deformation (rise) of the formation, which will lead to the decrease of the groundwater level. The deformation of the formation can also lead to abnormal noise. The leakage of underground high temperature and high pressure charged gases will inevitably lead to the change of underground and

ground temperature, humidity, electromagnetic field, air composition and air ion concentration. So it's possible to predict earthquakes by monitoring changes in those parameters.

Because earthquakes that are harmful to human beings are mainly shallow-source earthquakes whose epicenters are located in or near cities, and because shallow-source earthquakes occur in geological fault zones, we can drill some detection wells (because the ground surface air is too influenced by atmospheric changes, some wells should be used to detect variations in various parameters of the air in the well) in the fault zones in or near cities, and install relevant instruments in the wells to monitor the changes in the composition, temperature, humidity, air pressure, ion concentration, sound waves and electromagnetic field of the air in the wells over the years, and centralize the data into data analysis centers and use computers to analyze the data and predict earthquakes. The parameters to be monitored are as follows:

1) Monitor changes in air composition in the detection wells, in particular whether the concentration of nitrogen and carbon dioxide increases. The high temperature and high pressure charged gases gathered underground before the earthquake is mainly nitrogen and carbon dioxide (this can be verified by detecting the composition of the gases released by the volcanic eruptions)

2) Monitor the changes of air temperature and humidity in the detection wells, because the seismic gas that accumulates underground is hot and dry gas, the release of this gas will change the temperature and humidity of the air in the wells.

3) Monitor the change of the concentration of charged ions in the air in the detection wells. Because the seismic gas accumulated underground is high temperature and high pressure charged gases, the leakage of these gases will change the concentration of charged ions in the air in the wells.

4) Monitor sound waves from underground in the detection wells to see if any anomalies occur. Because the underground high temperature and high pressure charged gas gathering will inevitably make the underground rock bed deformation or fracture, resulting in sound waves.

5) Monitor the change of electromagnetic field in the detection wells. Because underground high temperature and high pressure charged gas accumulation and leakage will produce electromagnetic fields.

6) Monitor the change of air pressure in the detection wells. Because the leakage of underground high temperature and high pressure charged gas will change the pressure of the air in the sealed wells.

7) Monitor air flow in uncovered detection wells. Because the leakage of high temperature and high pressure charged gas underground will produce underground wind.

8) Monitor the deformation and tilt of the detection wells. Because the underground high temperature and high pressure charged gas accumulation will lead to formation rise and deformation, resulting in deformation or tilt of the detec-

tion wells.

9) Monitoring the changes in groundwater level in detection wells. Because the underground high temperature and high pressure charged gas accumulation will lead to formation rise and deformation, resulting in the decline of the well water level.

10) Monitor whether the ground as a whole is rising (monitored by satellite and ground equipments). Because the accumulation of underground high temperature and high pressure charged gas will lead to the rise of the formation.

3.2. The Elimination of Earthquakes and the Utilization of Seismic Energy

The focus of the earthquake elimination is to eliminate the shallow earthquakes with a focal depth of less than 30 km in cities especially in large and medium-sized cities. Since most earthquakes occur in the fault zones, it can be considered to drill some deep Wells (such as 10 - 30 km deep) in the fault zones to release underground gas all the year round. Because fractures in the fault zone are covered by strata and connected to the mantle below, high pressure bubbles from the mantle can enter and accumulate in these fractures by the magmatic buoyancy. Since the underground gas that accumulates in the fault zone cannot be released into the atmosphere in time, it will accumulate more and more and the pressure will be higher and higher. When the pressure exceeds the limit, the accumulated high pressure gas will explode (the accumulated high pressure gas will open up the formation and release into the atmosphere). This is the process of shallow earthquakes. So deep Wells connected to fractures in the fault zone to release underground gas all year round can prevent shallow earthquakes from forming.

When technology is highly advanced in the future, the controlled release of this underground high pressure gas can also be used to generate electricity.

3.3. The Conjecture of the Cause of the Ice-Age

The temperature of the earth's surface (including the atmosphere) is affected by the radiation power of the sun, the radiation power of the sun received by the earth, and the rate at which geothermal energy (carried by magma and gases separated from it) escapes from the earth's crust.

Because the solar system moves periodically around the Milky Way, changes in the speed of the solar system's motion in the Milky Way's etheric system affect the solar radiation power, the radius of the Sun and the Earth, the distance between the Sun and the Earth, and the rate at which the Earth's internal thermal materials (magma and gases isolated from it) spill from the Earth's crust. So the surface temperature of the earth changes periodically.

1) The increase in the solar system's speed of motion in the Milky Way's etheric system will lead to a decrease in solar radiation power: According to formula (5.17) [3], the increase of the sun's motion speed in the ether of the ga-

laxy will slow down time on the sun, that is, the speed of energy (ether) flowing from the sun to space will decrease, so the radiation power of the sun will decrease.

2) The increase in the solar system's speed of motion in the Milky Way's etheric system will lead to a reduction in the radius of the Sun and Earth: According to formula (5.31) [3], the increase in the speed of the solar system in the galactic ether will lead to a reduction in the radius of the atoms and ions that make up the sun and earth, thus resulting in a reduction in the radius of the sun and earth. A reduction in the radius of the sun and earth will result in a decrease in the power of solar radiation received by the earth.

3) The increase in the solar system's speed of motion in the Milky Way's etheric system will lead to a decrease in the distance between the Sun and Earth: It is not difficult to prove that an increase in the speed of the solar system in the galactic ether would also lead to a decrease in the distance between the sun and the earth. The decreasing distance between the sun and the earth will lead to the increase of solar radiation power received by the earth.

4) The accelerated motion of the solar system in the galactic ether will slow the rate at which hot material (magma and gas separated from the magma) in the earth's interior escape from the crust: According to formula (5.31) [3], when the solar system moves faster in the Milky Way, the speed at which the atoms that make up the earth expand with the expansion of the universe will decrease. The decrease in the rate of atomic expansion of underground magma will lead to the decrease in the rate of magmatic gas production, which will lead to the decrease in the rate of magma and high temperature gas flowing out of the earth's crust.

Since three of the above four factors, 1), 2) and 4), cause the Earth's temperature to drop, we have every reason to assume that the Earth's Ice-Age is caused by the acceleration of the motion of the solar system relative to the Milky Way at certain intervals. The greater the speed of the solar system relative to the Milky Way, the more the earth's surface temperature drops.

3.4. The Conjecture of the Contraction and Orogenesis of the Earth Lithosphere during the Ice-Age

Before the rupture of the earth's lithosphere, when the solar system moves to the regions of the Milky Way that make the solar system move faster, that is, those that put Earth into the Ice-Age, by the formula (5.31) [3], if the Earth's magma atoms begin to contract, it will cause the magma to stop producing gas, so the magma and the gas layer below the lithosphere will contract. In addition, the earth's surface temperature will decrease during the Ice-Age. The contraction of magma and gas beneath the lithosphere, coupled with the lowering of the earth's surface temperature, causes the lithosphere to contract. The contraction of the lithosphere causes the earth's surface to wrinkle. This may be another possible cause of orogeny. So it is possible that the orogeny was caused by the earth's

contraction and expansion. During the Ice-Age the earth contracted and the lithosphere contracted and wrinkled to build mountains. After the Ice-Age, the earth expanded and the lithosphere continued to build mountains under local uplift caused by underground high-pressure gas and expanding magmatism.

4. Conclusions

1) Cosmic expansion is the dynamic force source of all planetary geological tectonic movements;

2) As the universe expands, the underground magma atoms are expanding and releasing gases. The expanding magma and high-pressure gases gathered under the lithosphere are pushing the earth to expand, changing the lithosphere shape and leading to the continuous generation of volcanic eruptions and earthquakes.

3) Seafloor spreading and continental drift are caused by the expansion of magma inside the earth, which pushes the earth's expansion so that the plates move away from each other, the gaps between the plates keep widening and new oceanic crust is constantly formed in the gaps. It is the expanding forces of magma, not the mantle convection, that cause the seafloor to spread. The so-called mantle plume and plate subduction phenomenon do not exist;

4) Since the earth is expanding and the plates are moving away from each other, the phenomena of plate collision and extrusion do not exist. Therefore, the theory that earthquakes are caused by the release of the elastic potential energy accumulated by plate extrusion and collision is not true.

5) The earthquake is caused by the explosion of high temperature and high pressure gas formed by the accumulation of charged gas molecules separated by underground magma;

6) The Earth's Ice-Age may have been caused by the acceleration of the speed of movement of the solar system relative to the Milky Way as it moves into certain regions of the Milky Way. The greater the speed of the solar system relative to the Milky Way, the greater the drop in Earth surface temperature.

Conflicts of Interest

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

References

[1] Ma, Z.J., Du, P.S and Hong, H.J. (2003) structure and Dynamics of the Earth, In: *Chapter* 1, *Section* 1.6 *Hypothesizes on the Dynamic Force sources for the Tectonic Movement of the Earth*, 1st ed. Guangdong Science and Technology Press, Guangzhou.

[2] Wu, Q.J. (2001) *Recent Developments in World Seismology*, No. 8.

[3] Wang, J.A. (2019) *Journal of Modern Physics*, **10**, 1615-1644. https://doi.org/10.4236/jmp.2019.1014107

[4] Houlton, B.Z., Morford, S.L. and Dahlgren, R.A. (2018) *Science*, **360**, 58-62. https://doi.org/10.1126/science.aan4399

[5] Chen, Z.G. (1992) *Journal of Hebei University of Geosciences*, No. 6, 586-594.

[6] Wang, S.C. Sina Blog, http://blog.sina.com.cn/s/articlelist_1399906103_0_1.html

[7] Liang, G.H. (2016) *Encyclopedia Knowledge*, No. 14, 6-13.

[8] Yue, Z.Q. (2013) The Causes That Each Airs His Own Views for Strong Seismicity and the Prediction, Science Blog. http://blog.sciencenet.cn/home.php?mod=space&uid=240687&do=blog&id=682428

[9] Shu, L.F., Tian, X.R. and Li, H. (1998) *World Forestry Research*, No. 6, 42-48.

Journal of Modern Physics, 2020, 8, 432-447
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Physical Constants as Result of the Many Hypercubic Lattices of a Multidirectional Discrete Space

Christiaan T. de Groot

Retired, University of Applied Sciences Utrecht, Netherlands
Email: ctdegroot@live.nl

How to cite this paper: de Groot, C.T. (2020) Physical Constants as Result of the Many Hypercubic Lattices of a Multidirectional Discrete Space. *Journal of Modern Physics*, **11**, 432-447.
https://doi.org/10.4236/jmp.2020.113027

Received: February 13, 2020
Accepted: March 14, 2020
Published: March 17, 2020

Abstract

A multidirectional discrete space consists of numerous hypercubic lattices each of which contains one of the spatial directions. In such a space, several groups of lattices can be distinguished with a certain property. Each group is determined by the number of lattices it comprises, forming the characterizing numbers of the space. Using the specific properties of a multidirectional discrete space, it is shown that some of the characterizing numbers can be associated with a physical constant. The fine structure constant appears to be equal to the ratio of two of these numbers, which offers the possibility of calculating the series of smallest numerical values of these numbers. With these values, a reasoned estimate can be made of the upper limit of the smallest distance of the discrete space of approximately the Planck length.

Keywords

Hypercubic Lattice, Multidirectional Discrete Space, Characterizing Numbers, Fine Structure Constant, Physical Constants, Planck Length

1. Introduction

Once in a while there is speculation about the origin of the physical constants of nature. As can be seen in the review article by Barrow [1], some authors have noticed coincidences between these constants and numbers of cosmological proportions, others see similarities between the constants and some numbers obtained in a combinatorial manner. Wesson [2] has proposed to completely eliminate the physical constants through appropriately adjusting the quantities and parameters in the fundamental equations of physics.

Here the possibility is investigated that the physical constants originate from a

discrete space with many spatial directions. A multidirectional discrete space consists of a large but limited number of sparsely connected four-dimensional subspaces, each forming one of the spatial directions. The huge number of mutually equal subspaces offers possibilities not existing in the usual continuous space. The set of all subspaces consists of various subsets with a certain property. Because each subset consists of subspaces encompassing the entire space, the number of subspaces characterizing the subsets is timeless and valid throughout space. This property makes these characterizing numbers an ideal candidate to be related to the physical constants. For some physical constants, indeed there appears to be such a relation. This will be shown in two main steps.

In the first step to be found in Chapter 2, some required physical quantities are expressed in the units of the multidirectional discrete space. Only two units are needed, the smallest distance of a discrete space and the unit representing the diversity of subspaces in which a phenomenon is present.

In the second step of Chapter 3, it is shown that the different directions of a field in a multidirectional discrete space can be treated as independent entities. When this is combined with the results of the units' conversion, it is showed that the field in a multidirectional discrete space likely has a granular appearance. By using this possibility in combination with the many subspaces of a multidirectional space, it is shown that some of the characterizing numbers are proportional to a physical constant. In Chapter 4 it is shown, based on the huge number of subspaces, that the fine structure constant is equal to the ratio of two characterizing numbers, making the constant being an expression of the space structure. This ratio of two natural numbers, allows calculating the minimal values of the numerator and denominator from the empiric value of the fine structure constant. In Chapter 5, the minimum values of nearly all characterizing numbers are calculated in case an additional structural element of the discrete space applies.

The last chapter shows how to determine the largest numerical value of smallest distance of the discrete space with the found values of the characterizing numbers. The distance found appears to be of the same order of magnitude as the Planck length.

2. Expressing Some Usual Physical Quantities in the Discrete Units

A set of loosely connected hypercubic lattices (as given in [3]) is used here as an example of a multidirectional discrete space. Such a space is discrete at the level of the lattice and in the number of hypercubic lattices it consists of. The discreetness of space is expressed in the unit-distance $\Delta x\left(=\Delta t\right)$, being the minimal distance both in space and time. Due to the many lattices, an object is present in a subset of all possible lattices. To describe a particle field, an additional unit S is required representing the lattices in which the object is present.

A physical quantity expressed in obvious unit Δx and in the unusual unit S is

called a discrete quantity, indicated in the subscript: q_d. Curiously, no more units are needed. To incorporation of the physical quantities in a discrete space, the centimetre and the second need to be converted into the unit-distances Δx in Δt. This fairly simple conversion will be described first. For that purpose it is assumed that the arithmetically obtained classical electron radius r_e, the distance from which the electron field is present, is directly proportional to the smallest size L_{d0} of the electron field in a lattice: $r_e = L_{d0}\rho_{con}$, wherein L_{d0} is expressed in Δx.

In a lattice, the distances Δx and Δt are equal: $\Delta x = \Delta t$. For the sake of clarity the two units Δx and Δt will be indicated separately.

Name Symbol	Definition	Unit	
Conversion constant	$\rho_{con} = r_e / L_{d0}$	cm/Δx.	(1)
Discrete distance	$x_d = x/\rho_{con}$	Δx.	(2)

The value of classical electron radius r_e of Equation (1) is an empirical fact. L_{d0} is defined in paragraph 4.3. The expectation could be that L_{d0} equals the smallest distance Δx in a discrete space. However, in paragraph 6.2 where a method is given to determine the maximal numerical value of L_{d0}, it turns out to be enormous in terms of Δx. With a given L_{d0} and r_e, the constant ρ_{con} is defined.

To determine the discrete velocity, it is assumed that the maximal velocity $v_{d\max}$ runs along the space-time diagonal of each lattice, making $v_{d\max} = 1$ so that:

$$\text{Discrete velocity} \quad v_d = v/c \quad \Delta x/\Delta t \ (=1), \tag{3}$$

$$\text{Discrete time} \quad t_d = tc/\rho_{con} \quad \Delta t. \tag{4}$$

To describe the mass in a discrete space in terms of discrete units, use is made of the expression $r_e = e^2/m_e c^2$ in the CGS-system (see, for example, [4]). Subsequently, the quantities in this equation are replaced by discrete quantities, wherein the discrete speed of light equals 1. This result in the expression is $R_{de} = e_d^2/m_{de}$ with which the discrete electron mass becomes:

$$\text{Discrete electron mass} \quad m_{de} = e_d^2/R_{de} \quad S/\Delta x. \tag{5}$$

The discrete elementary charge e_d is defined in the next chapter by Equation (14). R_{de} is the size of the virtual portion of the discrete field representing its field energy as defined in paragraph 4.3. There it is shown that in a lattice R_{de} is not equal to the smallest size L_{d0} of the electron field. The unit S represents the lattices over which the particle field is spread.

As explained under Equation (9), the discrete and usual mass only differ by a proportionality constant making that the mass ratios are equal:

$$m_{dp}/m_{de} = m_p/m_e \tag{6}$$

Expressed in discrete quantities, the discrete force is: $f_d = \Delta m_d v_d / \Delta t_d$ and the discrete work is the product $W_d = f_d x_d$. Together with (2), (3), (4) and (6), the relation between discrete work W_d and conventional one W is:

$$\text{Discrete work} \quad W_d = W m_{de}/m_e c^2 \quad S\Delta x/\Delta t^2. \tag{7}$$

The Coulomb energy expressed in the Gaussian unit system is: $U_c = e^2/r$ [4]. Transcribing of the usual quantities in this relation by the discrete ones gives the discrete potential energy of the electron:

$$\text{Discrete Coulomb pot. } U_{dC} = e_d^2/r_d \quad S/\Delta x, \tag{8}$$

in which r_d is the discrete distance between two particles. Above equation implies that the discrete elementary charge e_d is expressed in the unit $\sqrt{S}$. Another relation for the discrete electrical potential U_{dC} can be obtained by inserting the usual $U_C = e^2/r$ in the relation (7) between discrete and conventional work. Equating this relation for U_{dC} with (8) gives a relation between the discrete distance r_d and r. By using relation (2), these distances can be eliminated, resulting in a second expression for the conversion constant ($\Delta x = \Delta t$):

$$\rho_{con} = m_{de}/m_e c^2 \quad e^2/e_d^2 \quad cm/\Delta x. \tag{9}$$

The definition of the conversion constant ρ_{con} is based on an electron. Alternatively, another charged fundamental particle could be used. Because ρ_{con} is the relationship between different kind of spaces independent of the particle used to define it, it follows from (9) that the particle independence of e/e_d also applies to m_{dp}/m_p.

According to the Gaussian unit system, the electric field strength is written as: $E = er/r^3$ [4], resulting in the analogue discrete electric field strength of an electron at distance r_d of the electron core:

$$\text{Discrete field strength } E_d = e_d \boldsymbol{e}_r/r_d^2 \quad \sqrt{S}/\Delta x^2, \tag{10}$$

with $\boldsymbol{e}_r$ the unit vector between the two electrons. The discrete electric field strength will be used in the next chapter to relate e_d with a natural number characterizing the discrete space.

Replace the quantities in the de Broglie momentum equation $mv = h/\lambda$ by the converted ones using Equations (2), (3), (6) and (9):

$$p_d = m_d v_d = hce_d^2/e^2\lambda_d \quad S/\Delta x.$$

By defining the discrete Planck's constant as $h_d = hce_d^2/e^2$, the discrete de Broglie momentum relation becomes:

$$\text{Discrete de Broglie } p_d = h_d/\lambda_d \quad S/\Delta x.$$

In $h_d = hce_d^2/e^2$, the usual Planck's constant h can be replaced by the fine-structure constant $\alpha = 2\pi e^2/hc$ written in the Gaussian unit system:

$$\text{Discrete Planck's const. } h_d = hce_d^2/e^2 = e_d^2(2\pi/\alpha) \quad S. \tag{11}$$

By applying the discrete de Broglie relation to a photon, the discrete momentum of the photon becomes:

$$\text{Discrete momentum } p_{d\,\text{photon}} = h_d/\lambda_{d\,\text{photon}} \quad S/\Delta x.$$

The conversion of the usual photon energy $W_{\text{photon}} = hc/\lambda_{d\,\text{photon}}$ into discrete quantities by using (2), (7), (9) and (11) gives

$$\text{Discrete energy } W_{d\,\text{photon}} = h_d/\lambda_{d\,\text{photon}} \quad S/\Delta x. \tag{12}$$

3. Discrete Physical Quantities Expressed in Characterizing Numbers

The correspondence of physical properties with space characteristics is possible in a multidirectional discrete space, a discrete space with many spatial directions within an overarching four-dimensional space. Reference [3] gives an example of such a space where the space consists of very many hypercubic lattices, the subspaces.

A multidirectional discrete space consists of an enormous number of only sparsely interconnected subspaces, each having a different spatial direction. The set of all subspaces can be subdivided into several sets containing a number of subspaces with a specific property, each number being one of the characterizing numbers of the discrete space. These are natural numbers of limited size which are valid throughout space, such as the total number of subspaces $n_{\text{subspaces}}$ or the number of subspaces with different spatial directions $n_{\text{directions}}$. It can easily be shown that $n_{\text{subspaces}} \gg n_{\text{directions}}$ [3]. A phenomenon placed in such a space is scattered over a subset of n_{particle} subspaces.

Using of the expressions for the discrete quantities of the previous chapter, it will be demonstrated that there is a relation between the characterizing numbers and some of the physical constants. Thereby it is assumed that a field with a certain direction is only present in those subspaces with an axis in the direction of the action.

3.1. Elementary Charge in a Multidirectional Discrete Space

Within a tangle of subspaces, there is always a subspace that has a connecting axis through two objects. The unit vector $\boldsymbol{e}_r$ of Equation (10) is such a connecting axis within one of the subspaces. The electric field in one of subspaces will be used to obtain the expression of the field strength in the set of subspaces with an axis parallel to the unit vector $\boldsymbol{e}_r$.

Rewrite Equation (10) to the field in one subspace. With one subspace, there is no need for the dimension $\sqrt{S}$. Further dimensional considerations indicate that the field strength in one subspace is only determined by $1/r_d^2$. When the field strength in one lattice is as simple as possible, it is only determined by the mutual distances. The consequence is that in one subspace $e_d = 1$, so that the field strength in one of subspaces with spatial axis $\boldsymbol{e}_r$ is:

$$E_{d\,sub} = \boldsymbol{e}_r / r_d^2 \quad 1/\Delta x^2. \tag{13}$$

Each subspace of the discrete space consists of a three spatial axis-directions. In the set of $n_{\text{directions}}$ subspaces, each axis direction is by definition present in n_{pivot} subspaces, a property resulting from one of the regularities in the structure of the space [3]. This means that of the $n_{\text{directions}}$ subspaces there are n_{pivot} subspaces with parallel axes, from which follows that the total field in a certain direction consists of n_{pivot} subspaces each of strength $\boldsymbol{e}_r / r_d^2$. With this, the total discrete electric field strength $\boldsymbol{E}_d$ in the direction $\boldsymbol{e}_r$ at a distance r_d of the electron core is:

$$E_d = n_{\text{pivot}} E_{d\,sub} = n_{\text{pivot}} \boldsymbol{e}_r / r_d^2 \quad \sqrt{S} / \Delta x^2 .$$

Comparing this expression with (10) gives that the discrete elementary charge e_d is equal to n_{pivot}. So in a discrete space, the elementary charge represents the number of parallel axes within a set of $n_{\text{directions}}$ subspaces:

$$\text{Discrete elementary charge} \quad e_d = n_{pivot} \quad \sqrt{S} . \tag{14}$$

The above equality defines the discrete elementary charge in a multidirectional discrete space assuming that n_{pivot} is a known quantity (see Table 1).

3.2. The Granulation of the Discrete Field

Take an electric field in the absence of another field. To abstract from the influence of any other field, the distance r_d in expression (13) is rewritten in L_d, being a certain variable determining the field strength in one subspace. With omission of the unit vector $\boldsymbol{e}_r$, the electrical field $E_{d\,sub}$ in one subspace is:

$$E_{d\,sub} = 1 / L_d^2 \quad 1/\Delta x^2 .$$

Apparently, L_d is a distance measure making that the field in one subspace is determined by a surface area L_d^2. With the above equation, the energy density $E_{d\,sub}^2$ the field in one subspace is:

$$E_{d\,sub}^2 = 1 / L_d^4 \quad 1/\Delta x^4 .$$

The relation for the energy density indicates that the discrete field can be divided into four-dimensional cubic-like entities of size L_d^4 of which the surface area L_d^2 forms a part. To obtain the field energy out of the energy density, a cube-shaped spatial part of the field of size L_d^3 is taken. To not be limited in possibilities, assume that within one subspace for unknown reasons the spatial part of the field has a repetition period in time T_d, implying that it is only partially present in time. With the gap in time, the spatial part of the field will only be present uninterrupted during the time L_d, and is not present at all in the remaining period T_d. If present, the energy density of the part of the field is $1/L_d^4$, for the rest of the time T_d there is no energy density at all. On average over the period T_d, the energy density is L_d / T_d part of $1/L_d^4$, making:

$$\text{energy density} = 1 / L_d^3 T_d$$

with this, the average discrete field energy $W_{d\,sub}$ in one subspace of a spatial part of the size L_d^3 is equal to:

$$W_{d\,sub} = 1 / T_d \quad 1/\Delta t .$$

3.3. The Field Energy of a Discrete Particle Field

A particle field is unlimited present with decreasing strength at increasing distance to the particle core. For a particle field consisting of concentric parts, each part with size L_{di} have an associated repetition distance T_{di}, where $1/T_{di}$ is the contribution to the energy of the field. The sum of these contributions results in the field energy $1/T_{d\,\text{particle}}$ of the particle. Due to the finite nature of the total field

energy of a particle, all various parts can be replaced by one virtual quantity $R_{d\,\text{particle}}$ with repetition distance $T_{d\,\text{particle}}$. Being cumulative quantities, both $T_{d\,\text{particle}}$ and associated $R_{d\,\text{particle}}$ are real numbers.

The particle field is present in at least $n_{\text{directions}}$ subspaces, being the minimal number of subspaces with all possible different directions, a small subset of the set of all subspaces. With $n_{\text{directions}}$ mainly independent subspaces, the total field energy of the particle is the cumulative of the individual contribution $1/T_{d\,\text{particle}}$ of each subspace[1]:

$$W_{d\ \text{particle}} = n_{\text{directions}} / T_{d\ \text{particle}} \quad S/\Delta t.$$

when, as usual, the total field energy is equated to the particle mass, the discrete particle mass becomes:

$$m_{d\ \text{particle}} = n_{\text{directions}} / T_{d\ \text{particle}} \quad S/\Delta t. \tag{15}$$

3.4. Relation Wavelength and Photon Energy

With regard to motion in a discrete space, there is still no satisfactory description thereof, see [5] or [6]. Only in the case of the motion of a photon in a lattice, some remarks can be made based on the probable granulation of the field and its movement along the space-time diagonal.

In a multidirectional discrete space, the photon field is distributed over n_{photon} subspaces, each with an axis in the direction of motion. Empirically, after annihilation of an electron-positron pair in photons, there is no remainder of the initial electron fields. In the process, the electron field within $n_{\text{directions}}$ subspaces with all possible directions is transformed into a photon field within n_{photon} subspaces with an axis in the direction of motion. To make this transition possible, a subspace with an arbitrary direction is somehow connected to a subspace with a preferred axis. The space-point of the multidirectional space, being a time series of subspaces as explained in 4.1, forms such a connection. Apparently a transformation of a part of the field in the one subspace to another field in another subspace can take place via the space-point in an unknown manner. This being the case, every subspace with random axes has a unique subspace with a preferred axis, so that $n_{\text{photon}} = n_{\text{directions}}$ and

$$W_{d\,\text{photon}} = n_{\text{directions}} / T_{d\,\text{photon}}\ .$$

The time length $T_{d\,\text{photon}}$ is equal to the spatial wavelength of the photon: At the moment of the conversion, each part of the electron field is transformed into a part of the photon field, making that the photon must also consists of parts. The photon field is moving with $v_{\text{photon}} = 1$ so its parts are present along the space-time diagonal of the subspace, meaning that the parts are moving stepwise with step size $L_{d\,\text{photon}} = T_{d\,\text{photon}}$ if the parts are contiguous.

Let there be a perfect synchronization of the presence in time of all parts of

[1]Note that in general a particle could be present in more, not less, than $n_{\text{directions}}$ subspaces. Is that the case, it is necessary because the discrete Planck's constant applies to all types of particles to split n_{particle} in subsets of $n_{\text{directions}}$ subspaces.

the photon field. This synchronization ensures that during each period $T_{d\,\text{photon}}$ the parts are simultaneously present at location $z = nT_{d\,\text{photon}} \,|\, (n = \text{integer})$ in all n_{photon} subspaces involved. Due to the relatively long-term presence at the location of each step, the photon had a greater chance of interaction on the spot. By the fixed mutual distances of those locations, $L_{d\,\text{photon}}$ can be interpreted as the discrete wavelength $\lambda_{d\,\text{photon}}$ of the previous chapter:

$$W_{d\,\text{photon}} = n_{\text{directions}} / \lambda_{d\,\text{photon}} \quad S/\Delta x.$$

By equating this expression with the same expression for the photon energy of the last chapter (12), the discrete Planck's constant becomes:

$$\text{Discr. Planck's constant} \quad h_d = n_{\text{directions}} \quad S. \tag{16}$$

3.5. Relation Discrete Electron Radius and Its Repetition Frequency

Equating the two expressions (11), (16) for the discrete Planck's constant results in $n_{\text{directions}} = e_d^2 (2\pi/\alpha)$. When this relation is subsequently inserted in the mass expression (15) one gets:

$$m_{d\,\text{electron}} = e_d^2 (2\pi/\alpha) / T_{de}.$$

The same discrete electron mass is also determined by (5): $m_{d\,\text{electron}} = e_d^2 / R_{de}$. R_{de} is, according to Section 3.3, the size of a virtual spatial part in which the total electron field is concentrated. In a multidirectional discrete space, the electron is thus determined by both a spatial size R_{de} and its repetition time T_{de}, related by:

$$T_{de} / R_{de} = 2\pi/\alpha. \tag{17}$$

Equation (17) will be used to relate the origin of the fine structure constant to the structure of the discrete space considered[2].

4. The Fine Structure Constant Arising from the Space Structure

In addition to the diversity of almost independent subspaces with a spatial direction, the multidirectional discrete space has characteristics resulting from the arrangement of subspaces along the time axes. This arrangement, necessary to provide for the plurality of directions, is the cause of the fine structure constant.

The discrete space studied here consists of many four-dimensional lattices linked together over the time axis of each lattice. The subspaces differ only in the direction of the three spatial axes. The resulting number of subspaces due to this space structure is countless. This number becomes less exuberant when there are regularities in the presence of subspaces. The connection via a space-point of two subspaces at regular space-time positions in both the one as the other subspace is one of these regularities when it applies to all subspaces.

[2]In the case there is no time gap in the presence of a part of the field in one subspace, the discrete electron mass would be: $m_{d\,\text{electron}} = e_d^2 (2\pi/\alpha) / R_{de}$. This combined with (5) results in $1 = 2\pi/\alpha$, a difficult to explain result. The time gap is therefore a necessity.

In case these regularities in space-time positions are such that the set of subspaces is limited, the space is a discrete space determined by the structural regularities. This makes the multidirectional discrete space a framework whose regularities have yet to be found. The only criteria for this are the properties of the reality and the phenomena occurring therein.

4.1. The Space-Point of the Multiple Directional Space

Let one of the regularities of space be such that each time axis consists of groups of n_{point} directly connected vertices, each of them belonging to a lattice with a different spatial direction. After n_{point} vertices, there is a connection with a vertex of the original lattice. By this way of organizing, every lattice of the space remains an ordinary hypercubic lattice with equal smallest distances in time Δt and space Δx. A space-point is an endless series of vertices connected in time. It consists of equal parts of n_{point} vertices, each of them belonging to a different lattice. The parts are repeated over time, such that after n_{point} vertices the vertex belongs to the same subspace.

One of the supposed regularities of the space is that n_{point} represents the same number for every space-point. Because of the succession of n_{point} intermediate lattices, the smallest distance of a lattice Δt can be subdivided into even smaller differential time steps δt with:

$$\Delta t = n_{\text{point}} \delta t.$$

The space thus formed consists of parallel time axes by definition, the three spatial axes perpendicular thereto are only in exceptional cases parallel. In **Figure A1** of the **Appendix**, a visualization is given of a limited series of coupled lattices with different spatial orientations.

4.2. The Imbalance in the Occurrence of Subspaces

To keep all possibilities open, presume that there is an imbalance in the number of subspaces with all possible spatial directions $n_{\text{directions}}$ in relation to the number of subspaces of each space-point: $n_{\text{directions}} \gg n_{\text{point}}$. The consequence of this imbalance is that the set of $n_{\text{directions}}$ subspaces is part of a bundle of space-points. Because all subspaces are interconnected via the space-points and the diverse lattices, many chains of space-time connections can be distinguished comprising all $n_{\text{directions}}$ subspaces. Within these chains, take the chain consisting of $t_{\text{directions}}$ consecutive equal time steps δt that comprises all $n_{\text{directions}}$ subspaces of the bundle. Then the ratio

$$t_{\text{directions}} / \Delta t = n_{\text{directions}} / n_{\text{point}}$$

represents the duration of the chain in question expressed in Δt. The first and last subspaces of the chain will have three parallel spatial axes, but both subspaces cannot be the same because the ratio will generally not be an integer, necessary to be the same subspace. It will appear that above ratio is equal to $(2\pi/\alpha)$.

4.3. The Fine Structure Constant and the Particle Field

According to paragraph 3.3, the particle field in each subspace consists of parts of size L_{di}, each part with a repetition distance T_{di}. During the time $T_{di} - L_{di}$ no part is present in one subspace. Because the parts must always be present somewhere, the parts are spread in time and space over the $n_{\text{directions}}$ subspaces. The scattered parts of the particle field are contiguous, meaning that when in one subspace the presence of a part stops after the time L_{di}, the presence of another part starts in a subsequent subspace. This implies that there must be a connection between the successive parts in the different subspaces in an unknown way. These connections form a chain over $n_{\text{directions}}$ consecutive subspaces with each $t = L_{di}$ a subsequent connection in another subspace. The total time distance of the chain is then $T_{di} = L_{di} n_{\text{directions}} / n_{\text{point}}$ expressed in Δt. To have the beginning and end of the chain in the same subspace, the size of each L_{di} is such that T_{di} is an integer. So the following relation between of two ratios of integers applies:

$$T_{di} / L_{di} = n_{\text{directions}} / n_{\text{point}} .$$

According to paragraph 3.3, particle field will consist of numerous field parts of various sizes, each part contributing $1/T_{di}$ to the total energy. Therewith the total energy per subspace becomes:

$$1/T_{d\,\text{particle}} = \sum 1/T_{di} = \sum 1/L_{di}\, n_{\text{point}} / n_{\text{directions}} .$$

$T_{d\,\text{particle}}$ can also be expressed in the smallest size L_{d0} of the particle field

$$1/T_{d\,\text{particle}} = c_p / T_{d0} ,$$

in which c_p is an easy to define particle dependent constant. The term $1/T_{d\,\text{particle}}$ determines the discrete particle mass (15). Define the associated virtual spatial part $R_{d\,\text{particle}}$ by $1/R_{d\,\text{particle}} = \sum 1/L_{di} = c_p / L_{d0}$, where c_p is the same constant as in the above relation. When applied to the electron, the ratio becomes:

$$T_{de} / R_{de} = n_{\text{directions}} / n_{\text{point}} = T_{d0} / L_{d0} . \tag{18}$$

The value of L_{d0} can be calculated as shown in paragraph 6.2, however, that does not apply to R_{de} which determines the discrete electron mass (5).

By combining (17) and (18), T_{de}/R_{de} can be eliminated resulting in a particle independent relation for the fine structure constant in a discrete space:

$$n_{\text{directions}} / n_{\text{point}} = 2\pi/\alpha . \tag{19}$$

Equation (19) expresses the fine structure constant in just two characterizing numbers showing that the natural constant seems to stem from a discrete character of the space structure.

5. Calculation of the Numerical Values of the Characterizing Numbers

The fine structure constant $1/(2\pi/\alpha)$ is a well known real number. Because the fine structure constant is a ratio of two integers (19), this enables to calculate the minimal values of $n_{\text{directions}}$ and n_{point}.

Instead of calculating these values directly, a regularity in the arrangement of

the subspaces of the space-point will be postulated. The chosen regularity remarkably gives the possibility to calculate almost all values of the characterizing numbers of the discrete space in the form of a broken series of natural numbers. It is used that the empirical value of $(2\pi/\alpha)$ only is fraction larger than 861.

5.1. Clustering of Subspaces of the Space-Point as Regularity of the Space Structure

Investigate the possibility that the subspaces of each space-point can be grouped into clusters of contiguous subspaces with a certain property. Each cluster covers an equal time interval:

$$\Delta t_{\text{cluster}} = n_{\text{cluster}} \quad \delta t.$$

wherein n_{cluster} is the number of subspaces of the cluster and δt is the differential time step of the space-point. Let a space-point consists of n_{groupP} clusters:

$$n_{\text{point}} = n_{\text{groupP}} n_{\text{cluster}} \quad S. \tag{20}$$

Let the $n_{\text{directions}}$ subspaces with different direction also be composed by a set of n_{cluster} subspaces, each cluster containing mutually different spatial directions. When the set consist of n_{groupD} clusters, the total number of subspaces is:

$$n_{\text{directions}} = n_{\text{groupD}} n_{\text{cluster}} \quad S, \tag{21}$$

with $n_{\text{groupD}} > n_{\text{groupP}}$, because $n_{\text{directions}} > n_{\text{point}}$.

The last inequality means that the clusters must be present in a bundle of space-points. The space-points of the bundle do not mutually differ in the number of subspaces but in the spatial directions of the subspaces. Inserting equations (20) and (21) in relation (19) gives:

$$2\pi/\alpha = n_{\text{directions}}/n_{\text{point}} = n_{\text{groupD}}/n_{\text{groupP}}. \tag{22}$$

5.2. The Characterizing Number Which Determines a Cluster of Subspaces

Let each cluster be determined by n_{spin} with $n_{\text{spin}} < n_{\text{pivot}}$. n_{pivot} is textually described near Equation (14). The characterizing number n_{spin} stands for the number of subspaces within the cluster with a certain property. In the Appendix an example is given in which n_{spin} represents the number of subspaces in the cluster with a parallel axis. In accordance with the Appendix, n_{cluster} is written as a quadratic function of n_{spin} with parameters n_{c1} and n_{c2}. These are such that n_{cluster} is a natural number:

$$n_{\text{cluster}} = n_{\text{spin}}^2 + 2n_{c2}n_{\text{spin}} + n_{c1}.$$

Due to the natural numbers that make up the characterizing numbers, equation with a ratio such as (22) can be split into two linked equations. With the extended number of equations thus obtained, nearly all characterizing numbers can be expressed in n_{cluster} and n_{spin} as will be shown.

5.3. Factorizing of $(2\pi/\alpha)$

Use of the property that the value of $(2\pi/\alpha)$ is very close to an integer. With this,

$(2\pi/\alpha)$ can be expressed in an integer value multiplied by the fraction of n_{cluster} and n_{spin}:

$$(2\pi/\alpha) = Trunc(2\pi/\alpha)n_{\text{cluster}}/n_{\text{spin}}^2 = Trunc(2\pi/\alpha)\left(1+2n_{c2}/n_{\text{spin}}+n_{c1}/n_{\text{spin}}^2\right). \quad (23)$$

Use this equation and (22) to eliminate $(2\pi/\alpha)$:

$$n_{\text{groupD}}/n_{\text{groupP}} = Trunc(2\pi/\alpha)n_{\text{cluster}}/n_{\text{spin}}^2.$$

The above relation only consists of quantities representing an integer. The relation can therefore be split into two linked relations, one determining the numerator and the other the denominator:

$$n_{\text{groupD}} = \rho Trunc(2\pi/\alpha)n_{\text{cluster}}, \quad (24)$$

$$n_{\text{groupP}} = \rho n_{\text{spin}}^2. \quad (25)$$

with which n_{groupD} and n_{groupP} are expressed in n_{spin}. The parameter ρ must be such that both n_{groupD} and n_{groupP} are natural numbers.

More relations can be obtained by equating the expressions (11) and (16) for the discrete Planck's constant: $n_{\text{directions}} = e_d^2(2\pi/\alpha)$. Inserting the expressions (14), (21) and (23) in this equation results in

$$n_{\text{groupD}}n_{\text{spin}}^2 = n_{\text{pivot}}^2 Trunc(2\pi/\alpha),$$

where the parameter ρ is gone by equating. Together with (24), this relation gives an expression of n_{pivot} in n_{spin}:

$$n_{\text{pivot}}^2 = n_{\text{cluster}}n_{\text{spin}}^2. \quad (26)$$

5.4. Relation Discrete Electric Charge and the Space-Point

After combining (14), (20), (25), and (26) one obtains:

$$n_{\text{point}} = n_{\text{pivot}}^2 = \rho n_{\text{spin}}^2 n_{\text{cluster}} \quad S, \quad (27)$$

$$e_d^2 = n_{\text{pivot}}^2 = n_{\text{point}} \quad S.$$

n_{cluster} is a natural number. When (27) is rewritten as $n_{\text{pivot}} = \sqrt{\rho}n_{\text{spin}}\sqrt{n_{\text{cluster}}}$, this implies that n_{cluster} is quadratic:

$$n_{\text{cluster}} = \left(n_{\text{spin}}+n_{c2}\right)^2, \quad (28)$$

and $n_{c1} = n_{c2}^2$.

Combining relation (27) with (14) and (28) gives a simple expression for the discrete electric charge e_d:

$$e_d = n_{\text{pivot}} = \sqrt{\rho}\left(n_{\text{spin}}+n_{c2}\right)n_{\text{spin}} \quad \sqrt{S}. \quad (29)$$

Above relation implies that the numerical value of $\sqrt{\rho}$ is such that $\sqrt{\rho}\left(n_{\text{spin}}+n_{c2}\right)n_{\text{spin}}$ is a natural number. Because ρ must also meet the conditions formulated in the text under (25), the most simple value of $\rho = 1$. This will be used for further calculations.

5.5. The Minimum Numerical Values of the Characterizing Numbers

The numerical value of $(2\pi/\alpha)$ is known within a certain measurement uncertainty.

Table 1. Some smallest characterizing numbers of discrete space.

n_{c2}	n_{spin}	$n_{cluster}$ (28)	n_{pivot} (29)	n_{point} (27)	$n_{directions}$ (21)(24)	$1/\alpha$ (23)	{Δ}
1	76,263	$5.8162\ 10^{9}$	$5.8160\ 10^{9}$	$3.3827\ 10^{19}$	$2.9126\ 10^{22}$	137.035 999 706	{27}
1	76,264	$5.8164\ 10^{9}$	$5.8163\ 10^{9}$	$3.3829\ 10^{19}$	$2.9128\ 10^{22}$	137.035 999 658	{−20}
1	76,265	$5.8165\ 10^{9}$	$5.8164\ 10^{9}$	$3.3831\ 10^{19}$	$2.9129\ 10^{22}$	137.035 999 611	{−68}
2	152,525	$2.3264\ 10^{10}$	$2.3264\ 10^{10}$	$5.4122\ 10^{20}$	$4.6600\ 10^{23}$	137.035 999 729	{50}
2	152,526	$2.3265\ 10^{10}$	$2.3265\ 10^{10}$	$5.4124\ 10^{20}$	$4.6602\ 10^{23}$	137.035 999 706	{27}
2	152,527	$2.3265\ 10^{10}$	$2.3265\ 10^{10}$	$5.4125\ 10^{20}$	$4.6603\ 10^{23}$	137.035 999 682	{3}
2	…	…					

By using a search procedure, it is easy to calculate the numerical values of n_{spin} out of Equation (23) via the inequality: $(2\pi/\alpha)_{-} <= (2\pi/\alpha) <= (2\pi/\alpha)_{+}$. The following empirical values are used: $(2\pi/\alpha)_{+-} = 861 * 1.00002622502(69)$.

The calculation results in a series of natural numbers for n_{spin} in dependence of n_{c2}. Table 1 shows only the smallest values for n_{spin} and, in addition, the related numerical values of some characterizing numbers. Of the numerical values shown in Table 1 are only those of n_{spin} the exact values. The other values are rounded numbers whose exact values can easily be calculated with the value of n_{spin} and the specified equations. To check, the last column shows the back-calculated fine structure constant using (23). For convenience: $1/\alpha = 137.035999679(94)$.

The individual values of the characterizing numbers have no error margins. The error margins can be found in the range of numerical values. Only the values of n_{spin} are contiguous, the other numbers have a broken range of values.

6. Concluding Remarks

An old question is whether the real space is a continuous or discrete. For a long time the possibility of a discrete space was not taken seriously, mainly because of the fundamental problems with such a space such as the motion problem and the anisotropy problem [6]. With the introduction of a multidirectional hypercubic lattice seems the anisotropy problem, the problem with the many directions, to have been tackled. The multitude of equal discrete subspaces therein offers the possibility that the physical constants arise from this omnipresent abundance. In a continuous space, these constants are only difficult to explain facts.

6.1. The Physical Constants as a Property of the Discrete Space

The multidirectional discrete space differs from the usual Euclidean space only at the level of the smallest distance Δx. Within Δx, the space has the counterintuitive property that there are no interconnections between space-points. Also the space is unusual in the division into many sparsely connected subspaces which differ only in the spatial direction. The scarce connections enable to treat the subspaces as almost completely separate items. This has repercussions for the

positioning of fields in a multidirectional discrete space.

The first consequence is that an additional unit is required indicating the subspaces in which the field is present. This makes it possible to convert the usual physical quantities into just two units. Secondly, the barely connected subspaces makes it very likely that the field is distributed over the various subspaces as loosely connected entities acting almost separately. This enables to express the discrete elementary charge in a number characterizing the discrete pace (14).

The implication of the units-conversion is that fields in a multidirectional discrete space should be described as a set of granulations. The graining allows expressing the discrete constant of Planck as a characterizing number of the multidirectional space (16). The consequence is that in each subspace these grains must have a gap in the presence over time, whereby the different granules form a consecutive series present in a set of subspaces. The distribution of parts of the particle field over subspaces belonging to multiple space-points can be seen as the reason for the occurrence of the fine structure constant.

6.2. The Maximal Value of the Smallest Distance of Discrete Space

By adding another regularity to the, as far as known, regularities in the space structure, nearly all characterizing numbers can be calculated from the fine structure constant. The diverse characterizing numbers are presented as a range of values starting with the minimal one. The size of these values depends on the measurement accuracy of the fine structure constant. The sheer size of the values found is an indication of their realism.

Not all characterizing numbers are linked here to a physical constant, such as the total number of subspaces $n_{\text{subspaces}}$. This number must be much larger than the number of subspaces with different directions $n_{\text{directions}} = 2.91 \times 10^{22}$ (**Table 1**), indicating that the physical constant associated with $n_{\text{subspaces}}$ will be very tiny. The gravitational constant is good candidate.

To obtain an indication of the smallest size L_{d0} of a particle field, ratio (18) is combined with (22):

$$T_{d0}/L_{d0} = n_{\text{directions}}/n_{\text{point}} = n_{\text{groupD}}/n_{\text{groupP}} \ .$$

Above relation indicates that various smallest sizes L_{d0} are possible, depending on the various ways the fraction $n_{\text{directions}}/n_{\text{point}}$ can be expressed in smaller numbers by reduction using the numerical values of $n_{\text{directions}}$ and n_{point}. When each thus obtained L_{d0} belongs to a fundamental particle, a series of such particles is possible varying from particles with relatively light mass for large L_{d0} to heavier particles with a small L_{d0}.

The smallest size L_{d0} of discrete electron field is needed to determine the conversion constant ρ_{con}. The electron probably has the relatively large value $L_{d\ 0} = n_{point}$, enabling heavier particles with smaller L_{d0}. With $n_{\text{point}} = 3.3827 \times 10^{19}$ from **Table 1** and the empirical electron radius $r_e = 2.82 \times 10^{-15}\text{m}$, the conversion constant ρ_{con} can be calculated with (1). This

ρ_{con} inserted in (2) gives as the largest possible value of the unit-distance of the lattice:

$$\Delta x = 8.33 \times 10^{-35} \text{ m},$$

based on the electron radius $L_{d0} = 3.38 \times 10^{19}$ Δx.

Δx is in accordance with the Planck length $L_P = K_l 4.05 \times 10^{-35} \text{m}$, with K_l a dimensionless constant, as obtained by dimensional analysis (see Meschini [7]). This is remarkable because the method to obtain Δx only depends on the value of the fine structure constant and a model of electron radius L_{d0}, without the use of the gravitational constant.

Expressed in Δx, the size of the discrete electron radius L_{d0} is enormous. This implies that within L_{d0} the field is uniformly constant, because there are no smaller parts of the field than L_{d0}. Therewith, the problem with the electron radius is not solved, but shifted to the problem of describing fields in a discrete space.

Conflicts of Interest

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

References

[1] Barrow, J.D. (1981) *Quarterly Journal of the Royal Astronomical Society*, **22**, 388-420.

[2] Wesson, P.S. (1992) *Space Science Reviews*, **59**, 365-406. https://doi.org/10.1007/BF00242090

[3] De Groot, C.T. (2017) *Journal of Modern Physics*, **8**, 1175-1189. http://www.scirp.org/journal/JMP/ https://doi.org/10.4236/jmp.2017.88078

[4] Landau, L. and Lifshitz, E. (1971) The Classical Theory of Fields. Oxford: Pergamon Press, 90, 91, 88.

[5] van Bendegem, J.P. (1995) *Logique et Analyse*, **38**, 127-150.

[6] Forrest, P. (1995) *Synthese*, **103**, 337, 339. https://doi.org/10.1007/BF01089732

[7] Meschini, D. (2007) *Foundations of Science*, **12**, 277-294. https://doi.org/10.1007/s10699-006-9102-3

Appendix A. A Possible Clustering of Subspaces

Take the possibility that the subspaces of a space-point are clustered into groups of subspaces with a spatial axis in a specific direction. The direction differs for each cluster. Within each cluster there are $n\|\|$ subspaces with mutual parallel spatial axes, the other axes of these subspaces have arbitrary directions. All clusters have an equal number of n_{cluster} subspaces.

Take as parameter the number of subspaces in a cluster with a parallel axis: n_{spin}. Let these subspaces be evenly spread over the time axis with an mutual time distance proportional to n_{spin}: $n_{\text{spin}} + n_c$ (see Figure A1). So

number of subspaces with parallel axes: $n\| = n_{\text{spin}}$,

time difference of the consecutive parallel axes: $\Delta t\| = n_{\text{spin}} + n_c$,

time interval to the next cluster Δt_c,

gives as the number of subspaces of a cluster:

$$n_{\text{cluster}} = \Delta t_{\text{cluster}} = n\|.\Delta t\| + \Delta t_c = n_{\text{spin}}^2 + n_c n_{\text{spin}} + \Delta t_c$$

Depicted a series of time connected vertices, each belonging to different subspaces, forming a small part of the time-axis of a space-point. A cluster is formed by n_{cluster} vertices in which there are $n\|$ parallel spatial axes positioned at regular time interval $\Delta t\|$. The other spatial axes are random. Also in other clusters are subspaces with an axis parallel to ||. Only within the cluster the parallel axes are at regular distances in the time.

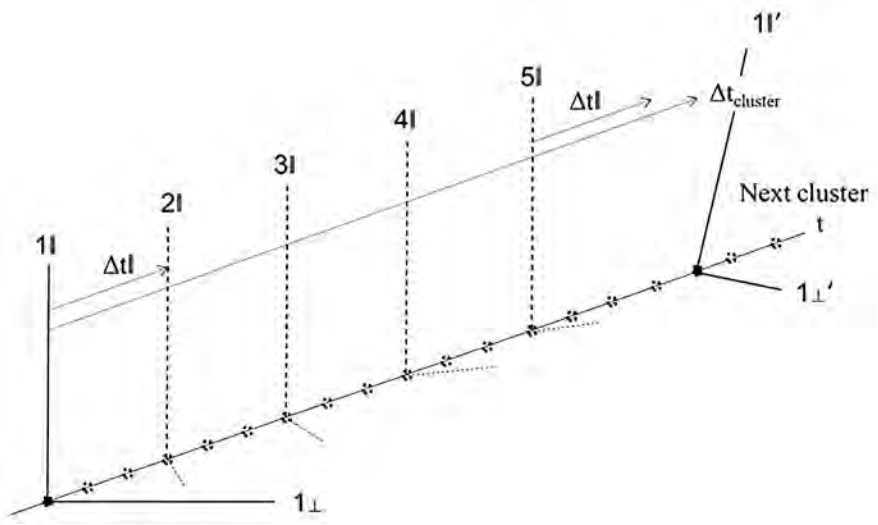


Figure A1. Clustering of subspaces.

Journal of Modern Physics, 2020, 11, 448-454
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Novel Model of Charged Leptons

Dianfu Wang, Xiao Liang, Yanqing Guo*

School of Science, Dalian Maritime University, Dalian, China
Email: wangdfu@dlmu.edu.cn, *yqguo@dlmu.edu.cn

How to cite this paper: Wang, D.F., Liang, X. and Guo, Y.Q. (2020) A Novel Model of Charged Leptons. *Journal of Modern Physics*, **11**, 448-454.
https://doi.org/10.4236/jmp.2020.113028

Received: February 21, 2020
Accepted: March 17, 2020
Published: March 20, 2020

Abstract

A novel model of charged leptons is presented, which contains two basics hypotheses. The first hypothesis is that the Yukawa coupling between Higgs field and charged leptons is the weak interaction, the Higgs field is a scalar intermediate boson which changes the chirality of charged leptons in the weak interaction. The other hypothesis is that the flavor eigenstates of charged leptons are the superposition states of left-handed and right-handed elementary Weyl spinors before the electroweak symmetry breaking. According to this model, the Yukawa coupling constants between Higgs field and three generations of charged leptons are considered to be a universal constant, and the difference of the masses of different charged leptons is due to the different left-right mixing angles of their flavor eigenstates.

Keywords

Charged Leptons, Weak Eigenstates, Left-Right Mixing Angle

1. Introduction

Up to now, with the discovery of the 125 GeV Higgs boson in 2012 at the CERN LHC [1] [2], the Weinberg-Salam (WS) model [3] [4] of the electroweak interaction of particle physics stands triumphant, and almost all relevant experimental results in particle physics are consistent with this model. Although the WS model has achieved impressive success in correlating all observed low-energy data in terms of a very few parameters, it cannot be called perfect. For example, the model is based on too many assumptions and leaves many fundamental questions unanswered. The success of the WS model only involves the gauge sector of the theory, in which only one free parameter, the Weinberg angle θ_W, is used to understand numerous neutral-current data. But for the fermionic sector, the fermion mass spectrum ranges from 170 GeV of the top-quark to 0.511×10^{-3} GeV of the electron. We do not know why there exists such a large difference

among the masses of these fermions, and we have no deep understanding of the origin of these masses. Since the fermion masses are related to the Yukawa couplings, we can only understand these differences and the mass-generation phenomenon by studying the Yukawa couplings in detail [5] [6] [7] [8].

2. A Charged Lepton Model

Quarks and leptons constitute the basic building blocks of matter in the standard model. In the WS model, there are three generations of quarks and leptons with identical quantum numbers but different masses. Since quarks are supposed to be confined by virtue of their strong interactions, the meaning of the quark mass is not obvious [9]. Therefore, in the present paper, for simplicity, we will limit ourselves to three generations of leptons. Let's start with discussing the lepton sector. Observationally, we must incorporate a neutral, left-handed neutrino (For simplicity, the neutrino is assumed to be massless, and hence right-handed neutrino is absent.) along with a charged lepton, which can be considered to be the sum of left-handed and right-handed Weyl spinors. The left-handed fermions form an isodoublet, consisting of the neutrino and the charged lepton:

$$L_\ell = \frac{1}{2}\left(1+\gamma_5\right)\begin{pmatrix}\nu_\ell\\ \ell\end{pmatrix} = \begin{pmatrix}\nu_\ell\\ \ell\end{pmatrix}_L, \tag{1}$$

while the right-handed sector consists of an isosinglet, the right-handed charged lepton:

$$R_\ell = \frac{1}{2}\left(1-\gamma_5\right)\ell = \ell_R, \tag{2}$$

where $\ell = e, \mu, \tau$ denote the flavor indexes of the charged leptons, and the subscripts L and R refer to the left- and right-handed components, respectively. In the standard WS model, the usual default setting is that the flavor eigenstates of charged leptons are the same as their mass eigenstates (An eigenstate of finite mass is a superposition of left and right-handed states with equal weight [9]), *i.e.*,

$$\ell = \ell_L + \ell_R. \tag{3}$$

with the chiral fermions given in Equation (1) and Equation (2), the most general $SU(2)_L \times U(1)_Y$ gauge-invariant lepton part of the Lagrangian density of WS model can be written as

$$\begin{aligned}\mathscr{L}_\ell = &-\bar{L}_\ell \gamma_\mu \left(\partial_\mu - ig\frac{1}{2}\tau^i A_\mu^i + ig'\frac{1}{2}B_\mu\right)L_\ell \\ &- \bar{R}_\ell \gamma_\mu \left(\partial_\mu + ig'B_\mu\right)R_\ell - G_\ell\left(\bar{L}_\ell \phi R_\ell + \bar{R}_\ell \phi^\dagger L_\ell\right).\end{aligned} \tag{4}$$

By defining the mass eigenstates of gauge fields

$$\begin{aligned}W_\mu^\pm &= \frac{1}{\sqrt{2}}\left(A_\mu^1 \mp iA_\mu^2\right),\\ Z_\mu &= \sin\theta_W B_\mu - \cos\theta_W A_\mu^3,\\ A_\mu &= \cos\theta_W B_\mu + \sin\theta_W A_\mu^3,\end{aligned} \tag{5}$$

where θ_W is the Weinberg angle with $\tan\theta_W = g'/g$, the interaction term that leptons coupled to gauge fields and Higgs field in Equation (4) can be written as

$$\begin{aligned}\mathscr{L}_\ell' = &+i\frac{g}{\sqrt{2}}\left[W_\mu^+\overline{(\nu_\ell)}_L\gamma_\mu\ell_L + W_\mu^-\overline{\ell}_L\gamma_\mu(\nu_\ell)_L\right]\\ &-i\sqrt{g^2+g'^2}Z_\mu\left[\frac{\overline{(\nu_\ell)}_L\gamma_\mu(\nu_\ell)_L - \overline{\ell}_L\gamma_\mu\ell_L}{2} + \sin^2\theta_W\overline{\ell}\gamma_\mu\ell\right]\\ &-i\frac{gg'}{\sqrt{g^2+g'^2}}A_\mu\overline{\ell}\gamma_\mu\ell - G_\ell\left(\overline{L}_\ell\phi R_\ell + \overline{R}_\ell\phi^\dagger L_\ell\right).\end{aligned} \tag{6}$$

The last term on the right side of Equation (6) is the gauge invariant Yukawa coupling between Higgs field and fermions,

$$\mathscr{L}_Y = -G_\ell\left(\overline{L}_\ell\phi R_\ell + \overline{R}_\ell\phi^\dagger L_\ell\right), \tag{7}$$

where G_ℓ, an additional parameter, gives the strength of the Yukawa coupling. In the standard WS model, all the charged leptons are massless as long as electroweak symmetry is unbroken, and they can get masses from Yukawa interactions only after electroweak symmetry breaking. A convenient way to implement this is to introduce a doublet Higgs field

$$\phi = \begin{pmatrix}\phi_+\\ \phi_0\end{pmatrix}, \tag{8}$$

where the subscripts refer to the electric charge. Replacing ϕ by its vacuum expectation value

$$\langle\phi\rangle = \begin{pmatrix}0\\ \upsilon\end{pmatrix}, \tag{9}$$

the mass term of the charged leptons can be given by the Yukawa coupling Equation (7) as

$$\mathscr{L}_{m_\ell} = -G_\ell\upsilon\overline{\ell}\ell. \tag{10}$$

Thus we can identify the masses of the charged leptons

$$m_\ell = G_\ell\upsilon. \tag{11}$$

However, it does not specify the value of the masses since the Yukawa coupling constant G_ℓ is arbitrary. So far, the standard model of the electroweak interaction has not given more further information on the origin of charged lepton masses. Furthermore, it is important to notice that since the strength of interaction between the Higgs field and any particle is proportional to the mass of that particle, this means that there is a new type of interaction that is different from the strong interaction, the electroweak interaction and even the gravitational interaction, which is very unnatural.

Since the mass of the Higgs boson is similar to that of W and Z bosons, so we could find it on the LHC in 2012. To some extent, they are all the same particles, or fields that presumably mediate weak interactions. In this sense, the Higgs boson will be no different from W and Z bosons. If these considerations are reasonable, the coupling constants of the four interaction terms in Equation (6) will

be of the same order of magnitude, which means the Yukawa coupling in Equation (4) can be written as

$$\mathscr{L}_Y = -g''\left(\bar{L}_\ell \phi R_\ell + \bar{R}_\ell \phi^\dagger L_\ell\right), \tag{12}$$

where g'' is the universal coupling constant of the Yukawa coupling. Based on the above discussion, we assume that the Yukawa coupling is the weak interaction, and the Higgs field is a scalar intermediate boson which changes the chirality of particles in the weak interaction. Since in low-energy processes the Higgs-meson and the W-meson carry small momentum, the propagators of them may be taken to be M_H^{-2} and M_W^{-2}. By considering Equation (6) and Equation (12), we can then obtain the following relation between the coupling constant g'' and the gauge coupling constant g as

$$\frac{g^2}{2M_W^2} = \frac{g''^2}{M_H^2}, \tag{13}$$

which gives the universal Yukawa coupling constant

$$g'' = \frac{g}{\sqrt{2}}\frac{M_H}{M_W} \tag{14}$$

Replacing ϕ by its vacuum expectation value $\langle\phi\rangle$, Equation (12) changes to be the mass term of the charged leptons

$$\mathscr{L}_{m_\ell} = -g''\upsilon\bar{\ell}\ell = -\frac{g\upsilon}{\sqrt{2}}\frac{M_H}{M_W}\bar{\ell}\ell, \tag{15}$$

The massess of the charged leptons derived from Equation (15) can be easily shown as

$$m_\ell = \frac{g\upsilon}{\sqrt{2}}\frac{M_H}{M_W} = \sqrt{2}M_H. \tag{16}$$

The comparison between Equation (16) and Equation (11) shows that this result is obviously wrong. The question now is whether the assumption Equation (12) is wrong or our understanding of it is incomplete. If we accept Equation (12), the only thing we can do is to reinterpret the flavor eigenstates of the charged leptons.

The form of the Lagrangian density given by Equation (4) indicates that the basic spinor fermion fields are not 4-component Dirac spinors, but rather their left and right-handed projections. Thus, for the weak interactions, the elementary entities are states of chirality with zero mass [9]. Based on these viewpoints, we propose a hypothesis here that the flavor eigenstates of the charged leptons are the superposition states of left-handed and right-handed Weyl spinors before electroweak symmetry breaking, *i.e.*,

$$\psi_\ell = \sqrt{2}\left(\cos\theta_\ell \ell_L + \sin\theta_\ell \ell_R\right). \tag{17}$$

where θ_ℓ are three left-right mixing angles (Hereafter, we will limit the angles θ_ℓ to $0 \le \theta_\ell \le \pi/2$.), and factor $\sqrt{2}$ is to ensure that Equation (17) is the same as Equation (3) when $\cos\theta_\ell = \sin\theta_\ell = 1/\sqrt{2}$. ℓ_L and ℓ_R refer to the

elementary entities, the left- and right-handed Weyl spinors, respectively.

By using Equation (17), the left- and right-handed charged leptons given in Equation (1) and Equation (2) should be replaced by

$$\begin{aligned} L'_\ell &= \frac{1}{2}\left(1+\gamma_5\right)\begin{pmatrix}\nu_\ell\\ \psi_\ell\end{pmatrix} = \begin{pmatrix}\nu_\ell\\ \psi_\ell\end{pmatrix}_L, \\ R'_\ell &= \frac{1}{2}\left(1-\gamma_5\right)\psi_\ell = \left(\psi_\ell\right)_R. \end{aligned} \tag{18}$$

Consequently, the Yukawa coupling in Equation (4) becomes

$$\mathscr{L}_Y = -\frac{g}{\sqrt{2}}\frac{M_H}{M_W}\left(\overline{L'_\ell}\phi R'_\ell + \overline{R'_\ell}\phi^\dagger L'_\ell\right). \tag{19}$$

By using Equation (18) and Equation (19), the Lagrangian density which characterizes the coupling of leptons to gauge fields and Higgs field in Equation (4) changes to be

$$\begin{aligned} \mathscr{L}'_\ell = &+i\frac{g}{\sqrt{2}}\left[W_\mu^+\overline{\left(\nu_\ell\right)}_L\gamma_\mu\left(\psi_\ell\right)_L + W_\mu^-\overline{\left(\psi_\ell\right)}_L\gamma_\mu\left(\nu_\ell\right)_L\right] \\ &-\frac{i}{2}\sqrt{g^2+g'^2}Z_\mu\left[\overline{\left(\nu_\ell\right)}_L\gamma_\mu\left(\nu_\ell\right)_L - \overline{\left(\psi_\ell\right)}_L\gamma_\mu\left(\psi_\ell\right)_L + 2\sin^2\theta_W\overline{\psi}_\ell\gamma_\mu\psi_\ell\right] \\ &-i\frac{gg'}{\sqrt{g^2+g'^2}}A_\mu\overline{\psi}_\ell\gamma_\mu\psi_\ell - \frac{g}{\sqrt{2}}\frac{M_H}{M_W}\left(\overline{L'_\ell}\phi R'_\ell + \overline{R'_\ell}\phi^\dagger L'_\ell\right). \end{aligned} \tag{20}$$

In Equation (20), all the terms except the Yukawa term do not involve cross term of left- and right-handed states. Thus, in these terms, according to quantum principles, it can be seen from Equation (17) that the left-handed fermion state $\left(\psi_\ell\right)_L = \sqrt{2}\cos\theta_\ell\ell_L$ and ℓ_L correspond to the same state, and for the same reason, the right-handed fermion state $\left(\psi_\ell\right)_R = \sqrt{2}\sin\theta_\ell\ell_R$ and ℓ_R correspond to the same state, too. However, in the Yukawa term, because of the interference between left-handed state $\left(\psi_\ell\right)_L$ and right-handed state $\left(\psi_\ell\right)_R$, the coefficients $\sqrt{2}\cos\theta_\ell$ and $\sqrt{2}\sin\theta_\ell$ are indispensable.

Based on the above considerations, after parameterization and unitary gauge transformation, replacing ϕ by

$$\begin{pmatrix}0\\ \upsilon+\eta\end{pmatrix}, \tag{21}$$

where η is the real Higgs field. The Yukawa term in Equation (20) changes to be

$$\mathscr{L}_Y = -\frac{g}{\sqrt{2}}\frac{M_H}{M_W}\sin\left(2\theta_\ell\right)\left(\upsilon+\eta\right)\overline{\ell}\ell. \tag{22}$$

Meanwhile, Equation (20) becomes

$$\begin{aligned} \mathscr{L}'_\ell = &+i\frac{g}{\sqrt{2}}\left[W_\mu^+\overline{\left(\nu_\ell\right)}_L\gamma_\mu\ell_L + W_\mu^-\overline{\ell}_L\gamma_\mu\left(\nu_\ell\right)_L\right] \\ &-\frac{i}{2}\sqrt{g^2+g'^2}Z_\mu\left[\overline{\left(\nu_\ell\right)}_L\gamma_\mu\left(\nu_\ell\right)_L - \overline{\ell}_L\gamma_\mu\ell_L + 2\sin^2\theta_W\overline{\ell}\gamma_\mu\ell\right] \\ &-i\frac{gg'}{\sqrt{g^2+g'^2}}A_\mu\overline{\ell}\gamma_\mu\ell - \frac{g}{\sqrt{2}}\frac{M_H}{M_W}\sin\left(2\theta_\ell\right)\left(\upsilon+\eta\right)\overline{\ell}\ell. \end{aligned} \tag{23}$$

This result is completely consistent with the standard WS model.

From Equation (22), we can obtain the masses of the charged leptons as

$$m_\ell = \sqrt{2} M_H \sin(2\theta_\ell). \tag{24}$$

Equation (24) shows that the difference of masses among the charged leptons of three generations comes from the different left-right mixing angles θ_ℓ of their flavor eigenstates.

Comparing Equation (11) with Equation (24), the usual Yukawa coupling constant G_ℓ is given by

$$G_\ell = \frac{\sqrt{2}}{\upsilon} M_H \sin(2\theta_\ell) = \frac{g}{\sqrt{2}} \frac{M_H}{M_W} \sin(2\theta_\ell). \tag{25}$$

From Equation (24), we can obtain

$$\sin(2\theta_\ell) = \frac{m_\ell}{\sqrt{2} M_H}, \tag{26}$$

where $m_\ell \ll \sqrt{2} M_H$ means $\theta_\ell \sim m_\ell / 2\sqrt{2} M_H \ll 1$. Now put $m_e \sim 0.511\,\text{MeV}$, $m_\mu \sim 105.66\,\text{MeV}$, $m_\tau \sim 1776.86\,\text{MeV}$ and $M_H \sim 125\,\text{GeV}$, one can obtain the magnitude of the three left-right mixing angles are about

$$\begin{aligned} \theta_e &\sim 1.45 \times 10^{-6}, \\ \theta_\mu &\sim 2.99 \times 10^{-4}, \\ \theta_\tau &\sim 5.03 \times 10^{-3}. \end{aligned} \tag{27}$$

This result implies that the charged leptons, as the superposition states of left- and right-handed elementary Weyl spinors, have an extreme left-right asymmetry. It is noteworthy here that the existence of left-right asymmetry should be related to the destruction of parity conservation in weak interactions [10].

3. Conclusion

The present paper mainly discusses the origin of the masses of charged leptons. The proposed model is based on two basic hypotheses: the first one is that there exists a universal constant of the Yukawa coupling term, this constant has the same order of magnitude as the gauge coupling constant; the other one is that the flavor eigenstates of the charged leptons are the superposition states of left-handed and right-handed Weyl spinors with different weights before electroweak symmetry breaking. Based on these two hypotheses, it is found that the difference in masses among the three generations of charged leptons is due to the difference in the left-right mixing angles in the definition of the flavor eigenstates of the different charged leptons. In fact, the lines of thought presented in this paper may be applied to three generations of neutrinos and even to the three generations of quarks. The origin of neutrino masses is one of the most compelling evidences for physics beyond the Standard Model (SM) [11], and there is no reliable theory to explain it. We will discuss the problem of neutrinos in our forthcoming works.

Acknowledgements

We thank the Editor and the referee for their comments.

Conflicts of Interest

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

References

[1] Chatrchyan, S., *et al.* (2012) *Physics Letters B*, **716**, 30-61.

[2] Aad, G., *et al.* (2012) *Physics Letters B*, **716**, 1-29.

[3] Weinberg, S. (1967) *Physical Review Letters*, **19**, 1264-1266. https://doi.org/10.1103/PhysRevLett.19.1264

[4] Salam, A. (1968) Elementary Particle Theory. In: Svartholm, N., Ed., Almquist and Forlag, Stockholm.

[5] Saldana-Salazar, U.J. (2016) *Physical Review D*, **93**, 013002. https://doi.org/10.1103/PhysRevD.93.013002

[6] Abdalgabar, A. (2017) arXiv:1706.02313[hep-ph]

[7] King, S.F. (2018) *Journal of High Energy Physics*, **9**, 069.

[8] de Varzielas, I.M. and King, S.F. (2019) *Physical Review D*, **99**, 095029.

[9] Huang, K.S. (2009) Quarks, Leptons & Gauge Fields. 2nd Edition, World Scientific Publishing Beijing Co Pte Ltd., Beijing.

[10] Lee, T.D. and Yang, C.N. (1956) *Physical Review*, **104**, 254-258. https://doi.org/10.1103/PhysRev.104.254

[11] Tanabashi, M., *et al.* (2018) *Physical Review D*, **98**, 030001.

Journal of Modern Physics, 2020, 11, 455-474
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

The Coherent State of the Landau Hamiltonian and the Relativistic Corrections to the Zeeman Effect in He^+ Ions

Kunnat Sebastian[1], Wangyao Li[2]

[1]Department of Physics and Applied Physics, University of Massachusetts Lowell, Lowell, MA
[2]School of Physical Science and Technology, ShanghaiTech University, Shanghai, China
Email: Kunnat_Sebastian@uml.edu, Liwy2@shanghaitech.edu.cn

How to cite this paper: Sebastian, K. and Li, W.Y. (2020) The Coherent State of the Landau Hamiltonian and the Relativistic Corrections to the Zeeman Effect in He^+ Ions. *Journal of Modern Physics*, **11**, 455-474.
https://doi.org/10.4236/jmp.2020.113029

Received: December 11, 2019
Accepted: March 17, 2020
Published: March 20, 2020

Abstract

We calculate the energy levels of He^+ ion placed in a uniform magnetic field directed perpendicular to the direction of its center of mass (CM) velocity vector, correct to relative order $\frac{v^2}{c^2}$. Our calculations are within the frame work of an approximately relativistic theory, correct to relative order $\frac{v^2}{c^2}$, of a two-particle composite system bound by electromagnetic forces, and written in terms of the position, momentum and spin operators of the constituent particles as proposed by Krajcik and Foldy, and also by Close and Osborn. Since the He^+ ion has a net electric charge, the total or the CM momentum is not conserved and a neat separation of the CM and the internal motion is not possible. What is new in our approach is that, for the basis states in a first order degenerate perturbation theory to calculate the effects of the external magnetic field, we use the direct product of the coherent state of the Landau Hamiltonian of the He^+ ion in a uniform magnetic field and of the simultaneous eigenstate of the internal Hamiltonian h, j^2, l^2, s^2 and j_z, where $\boldsymbol{j}$, $\boldsymbol{l}$ and $\boldsymbol{s}$ are the internal total, orbital and spin angular moments of the He^+ ion. The coherent state is an excellent approximation to the expected classical circular motion of the center of mass (CM) of the He^+ ion. In addition to the $Z^2\alpha^2$ corrections to the usual nonrelativistic results, including the small corrections due to the nuclear motion, we also obtain corrections which depend on the kinetic energy (E_{CM}) of the CM circular motion of the He^+ ion, in a nontrivial way. Even though these corrections are proportional to $\frac{E_{CM}}{Mc^2}$, where M is the mass of the He^+ ion, and are small for nonrelativistic CM motion, the results should be verifiable in careful experiments. Our results may also have application in astrophysical observations of the spectral

lines of He^+ ions in magnetized astrophysical objects.

Keywords

Relativistic Corrections, Zeman Effect, Helium Ion in a Magnetic Field, Coherent State, Landau Hamiltonian, Energy Levels

1. Introduction

When a composite system with a nonzero net electric charge is placed in a uniform magnetic field, neither the total canonical or mechanical momenta are conserved. So there is no inertial frame where the total momentum is zero at all times. So the problem of separating the c.m. motion from the internal motion for such a system will be different from that of an isolated system where the total or the c.m. momentum is conserved. In the past, several authors [1]-[8] have studied the problem of calculating the corrections to the internal energy levels of composite ions in a magnetic field. Some of them [1] [2] [3] [4] are based on a constant of motion introduced initially by Baye. And some of the works treated N-body problem [6] [8]. There are also aothors introducing new momentas using the coordinates transformation [5], however the coupling term is not small when one of the particles is much heavier than the other. In the paper [7], the the center of mass motion in an electromagnetic radiation is researched. In this paper, we take a different approach to this problem which is more physical. First of all we note that the overall or the c.m. motion of the ion in a uniform magnetic field under ordinary circumstances, is the familiar classical circular motion, if the uniform magnetic field is directed perpendicular to the plane of motion of the ion. In a recent paper [9] we have shown that the quantum state which most closely resembles the classical state of a charged particle moving in a uniform magnetic field, is the coherent state of the Landau quantum Hamiltonian of such a particle. In this paper, we will calculate the corrections to the energy levels of a H like ion (specifically He^+ ion) in a uniform magnetic field, by treating the state of the c.m. motion of the ion as a coherent state of the Landau Hamiltonian. One of the interesting features of our results is that the shifts in energy levels of the ion, including their first order relativistic corrections, depend on the energy of the c.m. motion, in a significant way. Even though these corrections are of the order of $\frac{E_{CM}}{Mc^2}$ times the usual results where M is the mass of the He^+ ion, and hence small for nonrelativistic c.m. motion, which we assume, it is important to include them, since in the future, more precise spectroscopic measurements may be able to detect such small corrections.

The format of the rest of the paper is as follows: In Section 2, we describe the approximately relativistic (correct to the order $\frac{v^2}{c^2}$) Hamiltonian of the isolated He^+ ion, and from there, how we obtain the Hamiltonian of the system in the

presence of an external magnetic field. We also introduce the approximately relativistic [10] [11] relations, again correct to order $\frac{v^2}{c^2}$, between the constituent and the c.m. variables. We then express these Hamiltonians in terms of these variables. In Section 3, we give the details of our perturbative calculation of the corrections to the energy levels due to the external uniform magnetic field. We also give particular attention to the basis states used in the first order perturbation calculations. The basis states we used, are the direct product of an internal state and the c.m. state which is the coherent state of the Landau Hamiltonian of a particle of charge $(Z-1)e$ and mass $M=(m_e+m_N)$ in a uniform magnetic field. In Section 4 we give our final results. Finally in Section 5, we make some concluding remarks.

2. Hamiltonian of the He^+ Ion in a Uniform Magnetic Field, Including Their First Order $\left(\frac{v^2}{c^2}\right)$ Relativistic Corrections

2.1. Isolated He^+ Ion

First let us consider the isolated He^+ ion, which is a weakly bound composite system of two particles where the first particle is the electron of mass m and electric charge -e, and spin $s=\frac{1}{2}$ and the second particle is the He^4 nucleus of electric charge Ze and mass m_N with zero spin and zero magnetic moment. We will put $Z=2$ only when we do numerical calculations. The Hamiltonian of such an electromagnetically bound system can be written, correct to order $\frac{v^2}{c^2}$ using the methods of Close and Osborn [12].

$$\begin{aligned} H=&\frac{p_1^2}{2m}+\frac{p_2^2}{2m_N}-\frac{p_1^4}{8m^3c^2}-\frac{p_2^4}{8m_N^3c^2}-\frac{Ze^2}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}+\frac{Ze^2}{4mm_Nc^2}\Bigg[\boldsymbol{p}_1\frac{1}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\cdot\boldsymbol{p}_2 \\ &+\boldsymbol{p}_1\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)\frac{1}{|\boldsymbol{r}_1-\boldsymbol{r}_2|^3}(\boldsymbol{r}_1-\boldsymbol{r}_2)\cdot\boldsymbol{p}_2+H.C.\Bigg]-\frac{Ze^2}{mm_Nc^2}\frac{\boldsymbol{s}_1\cdot\left[(\boldsymbol{r}_1-\boldsymbol{r}_2)\times\boldsymbol{p}_2\right]}{|\boldsymbol{r}_1-\boldsymbol{r}_2|^3} \\ &-\pi\frac{Ze^2}{4m^2c^2}\delta(\boldsymbol{r}_1-\boldsymbol{r}_2)+\frac{Ze^2}{2m^2c^2}\frac{\boldsymbol{s}_1\cdot\left[(\boldsymbol{r}_1-\boldsymbol{r}_2)\times\boldsymbol{p}_1\right]}{|\boldsymbol{r}_1-\boldsymbol{r}_2|^3} \end{aligned} \quad (1)$$

The first line in the above equation represents the kinetic energy terms, including their first order $\left(\frac{v^2}{c^2}\right)$ corrections and the Coulomb potential energy of the two charges Ze and -e. The second and third lines are the straight forward Breit interaction, resulting from the unretarded transverse one photon exchange. The fourth line represents the interaction between the spin magnetic moment of the electron and the magnetic field produced by the motion of the charged nucleus. The fifth line represents the effect of the zitterbewegung of the electron's motion, coming from the Darwin term, in the nonrelativistic reduction of the Dirac Hamiltonian of the electron. The sixth line represents the conventional

spin-orbit interaction, including the so called Thomas precession. The above expression for the Hamiltonian is also the same as the expression for the Hamiltonian of a two particle system interacting electromagnetically, given by Krajcik and Foldy [10], specialized to our case, where one of the particles He^4 nucleus has zero spin and zero magnetic moment. We have also put the electronic spin magnetic moment μ_e in their expressions [10] [11] as $\mu_e = -\frac{e}{2mc}$ which neglects the radiative corrections. This is consistent with the spirit of their approach which neglected the radiative corrections to the one photon exchange between two charged particles. It should also be noted we have used the Gaussian system of units, where as Close and Osborn [12] and Krajcik and Foldy [10] used rationalized Gaussian units. Also in Equation (1) we have used the system of units where $\hbar = 1$, but $c \neq 1$, where as the previous authors [10] [11] have put $\hbar = c = 1$.

It has already been demonstrated [10] [13] that the above Hamiltonian is part of a consistent relativistic theory to order $\frac{1}{c^2}$, by giving explicit expression [10] for the ten generators of the Poincare group in terms of the basic variables of the theory, $\boldsymbol{r}_\mu$, $\boldsymbol{p}_\mu$ and $\boldsymbol{s}_\mu$ $(\mu = 1,2)$ and showing that they satisfy the commutation relations of the Poincare group to the same $\frac{1}{c^2}$ order. The relativistic c.m. and internal variables are defined by the requirement that when the ten generators are expressed in terms of them, they take the single particle form [10] [11] [12] to the same order $\frac{1}{c^2}$.

Using this requirement, the relativistic relations between the constituent variables ($\boldsymbol{r}_\mu$, $\boldsymbol{p}_\mu$ and $\boldsymbol{s}_\mu$, $\mu = 1,2$) and the c.m. and the internal variables ($\boldsymbol{R}$, $\boldsymbol{P}$, $\boldsymbol{q}$, $\boldsymbol{p}$ and $\boldsymbol{\sigma}_\mu$, ($\mu = 1,2$)) for a two particle composite system made up of two particles, particle 1 being the electron of mass m and charge $-e$ and particle 2 being the nucleus of mass m_N and charge $+Ze$, are given by [10] [11] (correct to order $\frac{v^2}{c^2}$).

$$\begin{aligned}\boldsymbol{r}_1 = \boldsymbol{R} &+ \frac{m_N}{M}\boldsymbol{q} - \frac{1}{2}\frac{\boldsymbol{q}}{c^2}\left[\frac{m_N - m}{2mm_N M}p^2 + \frac{1}{2M^2}\boldsymbol{P}\cdot\boldsymbol{p}\right] + H.C. \\ &+ \frac{1}{2}\boldsymbol{q}\cdot\boldsymbol{P}\frac{1}{M^2c^2}\left(\frac{1}{2}\boldsymbol{p} - \frac{m_N}{m}\boldsymbol{p} - \frac{m_N}{2M}\boldsymbol{P}\right) + H.C. \\ &- \frac{1}{2mMc^2}\boldsymbol{p}\times\boldsymbol{\sigma}_e + \frac{m_N}{2mM^2c^2}\boldsymbol{P}\times\boldsymbol{\sigma}_e + \frac{Ze^2}{2M^2c^2}(m_N - m)\frac{\boldsymbol{q}}{q}\end{aligned} \tag{2a}$$

$$\begin{aligned}\boldsymbol{r}_2 = \boldsymbol{R} &- \frac{m}{M}\boldsymbol{q} - \frac{1}{2}\frac{\boldsymbol{q}}{c^2}\left[\frac{m_N - m}{2mm_N M}p^2 + \frac{1}{2M^2}\boldsymbol{P}\cdot\boldsymbol{p}\right] + H.C. \\ &+ \frac{1}{2}\boldsymbol{q}\cdot\boldsymbol{P}\frac{1}{M^2c^2}\left(\frac{1}{2}\boldsymbol{p} - \frac{m}{m_N}\boldsymbol{p} + \frac{m}{2M}\boldsymbol{P}\right) + H.C. \\ &- \frac{1}{2mMc^2}\boldsymbol{p}\times\boldsymbol{\sigma}_e - \frac{1}{2M^2c^2}\boldsymbol{P}\times\boldsymbol{\sigma}_e + \frac{Ze^2}{2M^2c^2}(m_N - m)\frac{\boldsymbol{q}}{q}\end{aligned} \tag{2b}$$

$$\boldsymbol{p}_1 = \boldsymbol{p} + \frac{m}{M}\boldsymbol{P} + \left[\left(\frac{m_N - m}{2mm_N}\right)p^2 + \frac{\boldsymbol{p}\cdot\boldsymbol{P}}{2M}\right]\frac{\boldsymbol{P}}{Mc^2} - \frac{Ze^2}{2M^2c^2}(m_N - m)\frac{1}{q}\boldsymbol{P} + \frac{Ze^2}{2M^2c^2}(m_N - m)(\boldsymbol{q}\cdot\boldsymbol{P})\frac{\boldsymbol{q}}{q^3} \tag{3a}$$

$$\boldsymbol{p}_2 = -\boldsymbol{p} + \frac{m_N}{M}\boldsymbol{P} - \left[\left(\frac{m_N - m}{2mm_N}\right)p^2 + \frac{\boldsymbol{p}\cdot\boldsymbol{P}}{2M}\right]\frac{\boldsymbol{P}}{Mc^2} + \frac{Ze^2}{2M^2c^2}(m_N - m)\frac{1}{q}\boldsymbol{P} + \frac{Ze^2}{2M^2c^2}(m_N - m)(\boldsymbol{q}\cdot\boldsymbol{P})\frac{\boldsymbol{q}}{q^3} \tag{3b}$$

$$\boldsymbol{s}_e = \boldsymbol{\sigma}_e - \frac{\boldsymbol{\sigma}_e \times (\boldsymbol{p}\times\boldsymbol{P})}{2mMc^2} \tag{4a}$$

$$\boldsymbol{s}_N = \boldsymbol{\sigma}_N + \frac{\boldsymbol{\sigma}_N \times (\boldsymbol{p}\times\boldsymbol{P})}{2m_NMc^2} \tag{4b}$$

where, $M = m + m_N$

$$\boldsymbol{q} = (\boldsymbol{r}_1 - \boldsymbol{r}_2)\big|_{\boldsymbol{P}=0} \tag{5}$$

$$\boldsymbol{p} = \boldsymbol{p}_1\big|_{\boldsymbol{P}=0} = -\boldsymbol{p}_2\big|_{\boldsymbol{P}=0} \tag{6}$$

$$\boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2 \tag{7}$$

$$\boldsymbol{R} = \frac{m\boldsymbol{r}_1 + m_N\boldsymbol{r}_2}{M}\bigg|_{\boldsymbol{P}=0} \tag{8}$$

We also notice, since we put $\hbar = 1$,

$$\left[q_i, p_j\right] = i\delta_{ij} \tag{9}$$

$$\left[R_i, P_j\right] = i\delta_{ij} \tag{10}$$

Also $\boldsymbol{R}$ and $\boldsymbol{P}$ commute with the internal variables $\boldsymbol{q}$ and $\boldsymbol{p}$ as well as $\boldsymbol{\sigma}_e$ and $\boldsymbol{\sigma}_N$. In our case, for the He^4 nucleus, $\boldsymbol{s}_N$ and $\boldsymbol{\sigma}_N$ both vanish. $\boldsymbol{\sigma}_N$ is the spin operator of the electron in the $\boldsymbol{P}=0$ frame. The reader can easily verify that when expressed in terms of the c.m. and the internal variables, the Hamiltonian of Equation (1) takes the single particle form,

$$H = \sqrt{c^2P^2 + h^2} \tag{11}$$

to order $\frac{v^2}{c^2}$. The other nine generators also take the single particle form again to order $\frac{v^2}{c^2}$, In particular, to order $\frac{v^2}{c^2}$, Equation (11) takes the form,

$$H = Mc^2 + h^{(0)}\left(1 - \frac{P^2}{2M^2c^2}\right) + h^{(1)} - \frac{P^4}{8M^3c^2} \tag{12}$$

where $h^{(0)}$ is the nonrelativistic internal Hamiltonian and $h^{(0)}$ is the first order $\left(\frac{v^2}{c^2}\right)$ correction to $h^{(0)}$. We find,

$$h^{(0)} = \frac{p^2}{2\mu} - \frac{Ze^2}{q} \tag{13}$$

$$
\begin{aligned}
h^{(1)} = & -\frac{p^4}{8m^3c^2} - \frac{p^4}{8m_N^3c^2} - \frac{Ze^2}{4mm_Nc^2}\left(\boldsymbol{p}\frac{1}{q}\cdot\boldsymbol{p} + H.C.\right) \\
& - \frac{Ze^2}{4mm_Nc^2}\left(\boldsymbol{p}\cdot\boldsymbol{q}\frac{1}{q^3}\cdot\boldsymbol{q}\cdot\boldsymbol{p} + H.C.\right) + \frac{Ze^2}{mm_Nc^2}\frac{\boldsymbol{\sigma}_e\cdot(\boldsymbol{q}\times\boldsymbol{p})}{q^3} \\
& - \pi\frac{Ze^2}{4m^2c^2}\delta(\boldsymbol{q}) + \frac{Ze^2}{2m^2c^2}\frac{\boldsymbol{\sigma}_e\cdot(\boldsymbol{q}\times\boldsymbol{p})}{q^3}
\end{aligned} \tag{14}
$$

In Equation (16) μ is the reduced mass

$$
\mu = \frac{mm_N}{m+m_N} = \frac{mm_N}{M} \tag{15}
$$

It should be stressed that Equation (1) would not have taken the single particle form of Equation (12) if we had used the nonrelativistic c.m. variables, which means neglecting the $\frac{1}{c^2}$ terms in Equation (2) and Equation (3). Also we will find with the nonrelativistic c.m. variables, the center of mass itself does not move uniformly and that the internal angular momentum, $\boldsymbol{l} = \boldsymbol{q}\times\boldsymbol{p}$ is not conserved. We should emphasize that $\boldsymbol{\sigma}_e$ in Equation (14) and elsewhere in the paper is not the Pauli matrix, but the spin $\frac{1}{2}$ operator of the electron, $\boldsymbol{s}_e$, in the $\boldsymbol{P}=0$ frame, which is really one half the usual Pauli matrix.

In order to derive Equations (12)-(14) we made use of the following convenient relations between the constituent and linearly independent c.m. and internal variables $\boldsymbol{R}$, $\boldsymbol{P}$, $\boldsymbol{q}$ and $\boldsymbol{p}$, which can be derived from Equations (2)-(4) using Equation (6) and Equation (7).

$$
\begin{aligned}
\boldsymbol{r}_1 - \boldsymbol{r}_2 = & \boldsymbol{q} - \frac{1}{2M^2c^2}\left(\frac{m_N}{m} - \frac{m}{m_N}\right)\left[(\boldsymbol{q}\cdot\boldsymbol{P})\boldsymbol{p} + \boldsymbol{p}(\boldsymbol{q}\cdot\boldsymbol{P})\right] \\
& - \frac{1}{2M^2c^2}(\boldsymbol{q}\cdot\boldsymbol{P})\boldsymbol{P} - \frac{\boldsymbol{\sigma}_e\times\boldsymbol{P}}{2mMc^2}
\end{aligned} \tag{16}
$$

$$
\begin{aligned}
|\boldsymbol{r}_1 - \boldsymbol{r}_2| = q\Bigg[& 1 - \frac{1}{2M^2c^2}\left(\frac{m_N}{m} - \frac{m}{m_N}\right)\times\frac{1}{q^2}\left[(\boldsymbol{q}\cdot\boldsymbol{P})(\boldsymbol{q}\cdot\boldsymbol{p}) + (\boldsymbol{q}\cdot\boldsymbol{p})(\boldsymbol{q}\cdot\boldsymbol{P})\right] \\
& - \frac{1}{2M^2c^2}\frac{1}{q^2}(\boldsymbol{q}\cdot\boldsymbol{P})^2 - \frac{\boldsymbol{q}\cdot(\boldsymbol{\sigma}_e\times\boldsymbol{P})}{2mMc^2q^2}\Bigg]
\end{aligned} \tag{17}
$$

The nonrelativistic internal Hamiltonian $h^{(0)}$ of Equation (13) gives the Bohr energy levels of a particle of reduced mass μ given by Equation (15). The first order relativistic correction to this Hamiltonian, $h^{(1)}$ of Equation (14), gives the fine structure corrections to the energy level due to the spin-orbit, magnetic and Darwin terms. It is important to note our expressions for $h^{(0)}$ and $h^{(1)}$ includes the effects of nuclear motion, that is terms of all orders in $\left(\frac{m}{m_N}\right)$ correct to order $\frac{v^2}{c^2}$. We will calculate the corrections to the Bohr energy levels due to $h^{(1)}$ in the first order degenerate perturbation theory where the basis states are chosen as the simultaneous eigen stats of $h^{(0)}$, l^2, σ_e^2, j^2 and

j_z, where,

$$\boldsymbol{l} = \boldsymbol{q} \times \boldsymbol{p} \tag{18}$$

$$\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{\sigma}_e \tag{19}$$

2.2. The Hamiltonian of the He+ Ion in a Uniform Magnetic Field

The main goal of this paper is to calculate the corrections to the eigenvalues or energy levels of $h^{(0)} + h^{(1)}$ of the He^+ ion due to the presence of an external uniform magnetic field, the so called Zeeman effect on a composite system with a net electric charge. The Hamiltonian of He^+ ion in the presence of an external magnetic field $\boldsymbol{B}$ is obtained from the Hamiltonian of Equation (1) by means of two requirements. 1) The resulting Hamiltonian should be gauge invariant in the sense that the Hamiltonian should be invariant under the gauge transformations

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \chi \tag{20}$$

where χ is an arbitrary function of $\boldsymbol{x}$ and $\boldsymbol{t}$, and A_μ is the four vector potential representing the external e.m. field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$. In our case where there is only an external magnetic field, this requirement reduce to,

$$\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \nabla \chi \tag{21}$$

2) The resulting Hamiltonian should reduce to the sum of the Foldy-Wouthuysen reduced Hamiltonians (to order $\frac{v^2}{c^2}$) of two relativistic free particles in the absence of any internal interaction between the particles.

The first requirement can be satisfied if every momentum operator $\boldsymbol{p}_\mu$ in Equation (1) is replaced by $\left(\boldsymbol{p}_\mu - \frac{e_\mu}{c}\boldsymbol{A}_\mu\right)$ where $\mu = 1,2$ is the particle label and $\boldsymbol{A}_\mu = \boldsymbol{A}\left(\boldsymbol{r}_\mu\right)$. The second requirement can be satisfied by adding two terms which depend on $\boldsymbol{B}$ and the spin operator of the electron and which are obtained from the Dirac Hamiltonian of the electron in the presence of a magnetic field by the Foldy-Wouthuysen reduction to order $\frac{v^2}{c^2}$. Using the two requirements, the Hamiltonian of the He^+ ion in the presence of an external uniform magnetic field to relative order $\frac{v^2}{c^2}$, is given by the following expression:

$$H = \frac{\left(\boldsymbol{p}_1 + \frac{e}{c}\boldsymbol{A}_1\right)^2}{2m} + \frac{\left(\boldsymbol{p}_2 - \frac{Ze}{c}\boldsymbol{A}_2\right)^2}{2m_N} - \frac{Ze^2}{\left|\boldsymbol{r}_1 - \boldsymbol{r}_2\right|} - \frac{\left(\boldsymbol{p}_1 + \frac{e}{c}\boldsymbol{A}_1\right)^2\left(\boldsymbol{p}_1 + \frac{e}{c}\boldsymbol{A}_1\right)^2}{8m^3c^2}$$
$$- \frac{\left(\boldsymbol{p}_2 - \frac{Ze}{c}\boldsymbol{A}_2\right)^2\left(\boldsymbol{p}_2 - \frac{Ze}{c}\boldsymbol{A}_2\right)^2}{8m_N^3c^2} + \frac{Ze^2}{4mm_Nc^2}\left[\left(\boldsymbol{p}_1 + \frac{e}{c}\boldsymbol{A}_1\right)\frac{1}{\left|\boldsymbol{r}_1 - \boldsymbol{r}_2\right|}\left(\boldsymbol{p}_2 - \frac{Ze}{c}\boldsymbol{A}_2\right)\right.$$
$$\left.\left(\boldsymbol{p}_1 + \frac{e}{c}\boldsymbol{A}_1\right)\cdot\left(\boldsymbol{r}_1 - \boldsymbol{r}_2\right)\frac{1}{\left|\boldsymbol{r}_1 - \boldsymbol{r}_2\right|^3} \times \left(\boldsymbol{r}_1 - \boldsymbol{r}_2\right)\cdot\left(\boldsymbol{p}_2 - \frac{Ze}{c}\boldsymbol{A}_2\right) + H.C.\right]$$

$$-\frac{Ze^2}{mm_N c^2}\frac{s_1\cdot\left[\left(\boldsymbol{r}_1-\boldsymbol{r}_2\right)\times\left(\boldsymbol{p}_2-\frac{Ze}{c}\boldsymbol{A}_2\right)\right]}{\left|\boldsymbol{r}_1-\boldsymbol{r}_2\right|^3}-\pi\frac{Ze^2}{4m^2c^2}\delta\left(\boldsymbol{r}_1-\boldsymbol{r}_2\right)$$
$$+\frac{Ze^2}{2m^2c^2}\frac{s_1\cdot\left[\left(\boldsymbol{r}_1-\boldsymbol{r}_2\right)\times\left(\boldsymbol{p}_1+\frac{e}{c}\boldsymbol{A}_1\right)\right]}{\left|\boldsymbol{r}_1-\boldsymbol{r}_2\right|^3}+\frac{e}{mc}\boldsymbol{s}_1\cdot\boldsymbol{B}-\frac{e}{4m^3c^3}\left[p_1^2,\boldsymbol{s}_1\cdot\boldsymbol{B}\right]_+ \tag{22}$$

The last two terms on the right hand side of Equation (25) represents the terms obtained from the second requirement mentioned above. In Equation (25) we have to keep terms up to order $\frac{1}{c^3}$ because of the $\frac{1}{c}$ factor in $\frac{1}{c}\boldsymbol{A}$ when $\boldsymbol{p}$ is replaced by $\left(\boldsymbol{p}-\frac{q}{c}\boldsymbol{A}\right)$ where q is the electric charge of the particle.

We have reasons to believe we are on the right track with the Hamiltonian of Equation (21). In previous works, [14] we have shown that the Hamiltonian in the presence of an external radiation field obtained by means requirements (1) and (2) reproduced the relativistic invariance of the one and the two photon transition amplitudes together with the use of the relativistic c.m. variables introduced above. It has also been shown [13] [14] [15] that the same Hamiltonian would reproduce the well-known low energy theorems on Compton scattering of photons by bound composite systems.

In Equation (22),

$$\boldsymbol{A}_\mu=\boldsymbol{A}\left(\boldsymbol{r}_\mu\right) \tag{23}$$

where $\mu=1,2$ is the particle label. In the case of the external uniform magnetic field $\boldsymbol{B}$,

$$\boldsymbol{A}_\mu=\frac{1}{2}\left(\boldsymbol{r}_\mu\times\boldsymbol{B}\right) \tag{24}$$

where we have chosen the symmetric gauge to define the vector potential. If $\boldsymbol{B}$ is along the Z-axis, the vector potential $\boldsymbol{A}$ will only have x and y components, so that

$$\boldsymbol{A}_\mu=-\frac{1}{2}By_\mu\hat{\boldsymbol{i}}+\frac{1}{2}Bx_\mu\hat{\boldsymbol{j}} \tag{25}$$

Substituting Equation (25) in Equation (22) and using the first order relativistic relations between the constituent and c.m. variables of Equations (2)-(4) and Equation (6) and Equation (7) in the resulting Equation (22) and after a series of simplifications and rearranging, we obtain the following expression for the Hamiltonian of He^+ ion in a uniform magnetic field directed along the Z-axis.

$$H=\frac{p^2}{2\mu}-\frac{Ze^2}{q}-p^4\left(\frac{1}{8m^3c^2}+\frac{1}{8m_N^3c^2}\right)+\frac{Ze^2}{2m^2c^2}\frac{1}{q^3}\boldsymbol{\sigma}_e\cdot\boldsymbol{l}\left(1+2\frac{m}{m_N}\right)$$
$$-\frac{Ze^2}{2mm_Nc^2}\left(\boldsymbol{p}\frac{1}{q}\cdot\boldsymbol{p}+\boldsymbol{p}\cdot\boldsymbol{q}\frac{1}{q}\boldsymbol{q}\cdot\boldsymbol{p}\right)-\pi\frac{Ze^2}{4m^2c^2}\delta\left(\boldsymbol{q}\right)$$

$$
\begin{aligned}
&+\left(\frac{P_X^2}{2M}+\frac{(Z-1)^2 e^2}{8Mc^2}B^2X^2\right)+\left(\frac{P_Y^2}{2M}+\frac{(Z-1)^2 e^2}{8Mc^2}B^2Y^2\right)\\
&-\frac{(Z-1)eB}{2Mc}L_Z+\frac{e^2B^2}{8c^2}\left(\frac{1}{m}+\frac{Z^2}{m_N}-\frac{(Z-1)^2}{M}\right)\left(X^2+Y^2\right)\\
&+\frac{(Z-1)eB}{2Mc}L_Z\frac{P^2}{2M^2c^2}-\frac{P^4}{8M^3c^2}+\frac{e^2B^2}{8mc^2}\left(q^2-q_z^2\right)\left(\frac{m_N^2}{M^2}+Z^2\frac{m}{m_N}\frac{m^2}{M^2}\right)\\
&+\left(\frac{p^2}{2\mu}-\frac{Ze^2}{q}\right)\frac{P^2}{2M^2c^2}\left(1-2\frac{m}{M}\right)-\frac{p^2}{2\mu}\frac{m}{M}\frac{P^2}{M^2c^2}\\
&-\frac{p_z^2}{2\mu}\frac{P^2}{2M^2c^2}+\frac{eB}{2mc}\left(\frac{m_N}{M}-Z\frac{m^2}{m_NM}\right)l_z+\frac{eB}{mc}\sigma_{ez}\\
&-\frac{eB}{2mc}\frac{p^2}{2m}l_z\left(\frac{m_N}{M}\frac{1}{mc^2}+\frac{m_N-m}{m_N}\frac{1}{Mc^2}\right)\\
&-\frac{ZeB}{2m_Nc}\frac{p^2}{2\mu}\left(\frac{m_N}{M}-\frac{m}{m_N}\right)l_z\frac{1}{Mc^2}\\
&-\frac{eB}{2mc}l_z\frac{P^2}{4M^2c^2}\left(1-3\frac{m}{M}-Z\frac{m}{m_N}\left(\frac{3}{2}+\frac{m}{2M}\right)\right)\\
&-\frac{eB}{2mc}\frac{Ze^2}{q}\frac{l_z}{2Mc^2}\left(1-\frac{m_N-m}{M}-Z\frac{m}{m_N}\left(\frac{3}{2}+\frac{m}{2M}\right)\right)\\
&-\frac{eB}{2mc}\frac{p^2-p_z^2}{2m}\frac{\sigma_{ez}}{2Mc^2}\left(1+Z\frac{m}{m_N}\right)-\frac{ZeB}{2m_Nc}\frac{Ze^2}{q^3}\left(q^2-q_z^2\right)\frac{\sigma_{ez}}{Mc^2}\\
&-\frac{eB}{2mc}\sigma_{ez}\frac{P^2}{2M^2c^2}\left(1+\frac{m}{M}-Z\frac{m}{M}\right)-\frac{eB}{2mc}\frac{p^2}{2m}\frac{L_Z}{Mc^2}\left(1-Z\frac{m^2}{m_N^2}\right)\\
&+\frac{eB}{2mc}\frac{Ze^2}{q}\frac{L_Z}{2Mc^2}\left(1-Z\frac{m}{m_N}\right)
\end{aligned}
\tag{26}
$$

where the c.m. orbital angular momentum operator,

$$\boldsymbol{L}=\boldsymbol{R}\times\boldsymbol{P}, \tag{27}$$

the operator $\boldsymbol{\sigma}_e$ is the spin operator of the electron in the $\boldsymbol{P}=0$ frame, except for the factor $\hbar$ which we have put equal to 1 in all of the above equations.

Several comments are in order about Equation (26). First, we have included terms of order B^2 only in the nonrelativistic part of the Hamiltonian. In the relativistic correction terms, we have kept only terms which are linear in *B*, mainly because the contribution of the quadratic and the higher order terms in *B* are demonstrably negligible, even for any reasonably strong values of the magnetic field. In the next section we will write the Hamiltonian of Equation (26) as the sum of an unperturbed Hamiltonian H_0 and a perturbation Hamiltonian *V*. The unperturbed Hamiltonian will contain $h^{(0)}+h^{(1)}$ plus the Landau Hamiltonian of a charged particle of electric charge $(Z-1)e$ and of mass *M*, moving in the xy plane perpendicular to the magnetic field $\boldsymbol{B}$ directed along the Z-axis. The remaining terms on the right hand side of Equation (26) will be treated as the perturbation *V*.

3. Perturbative Calculation of Corrections to the Energy Levels of He^+ Ion Due to the External Uniform Magnetic Field

For the first order perturbation calculation, we first write down our explicit expressions for H_0 and V.

$$\begin{aligned} H_0 = & \left(\frac{p^2}{2\mu}-\frac{Ze^2}{q}\right)-p^4\left(\frac{1}{8m^3c^2}+\frac{1}{8m_N^3c^2}\right) \\ & +\frac{Ze^2}{2m^2c^2}\frac{1}{q^3}\boldsymbol{\sigma}_e\cdot\boldsymbol{l}+\frac{Ze^2}{mm_Nc^2}\frac{1}{q^3}\boldsymbol{\sigma}_e\cdot\boldsymbol{l}-\pi\frac{Ze^2}{4m^2c^2}\delta(\boldsymbol{q}) \\ & -\frac{Ze^2}{2mm_Nc^2}\left(\boldsymbol{p}\frac{1}{q}\cdot\boldsymbol{p}+\boldsymbol{p}\cdot\boldsymbol{q}\frac{1}{q}\boldsymbol{q}\cdot\boldsymbol{p}\right)+\left(\frac{P_X^2}{2M}+\frac{(Z-1)^2e^2}{8Mc^2}B^2X^2\right) \\ & +\left(\frac{P_Y^2}{2M}+\frac{(Z-1)^2e^2}{8Mc^2}B^2Y^2\right)-\frac{(Z-1)eB}{2Mc}L_Z \end{aligned} \tag{28}$$

We can write H_0 as,

$$H_0 = h^{(0)}+h^{(1)}+H_0^L \tag{29}$$

where $h^{(0)}$ is the nonrelativistic internal Hamiltonian and $h^{(0)}$ is its first order (of order $\frac{v^2}{c^2}$) correction, proportional to $\frac{1}{c^2}$. H_0^L is the Landau Hamiltonian of a particle of mass M and electric charge $(Z-1)e$.

$$H_0^L = \frac{\left(P_X^2+P_Y^2\right)}{2M}+\frac{(Z-1)^2e^2B^2}{8Mc^2}\left(X^2+Y^2\right)-\frac{(Z-1)eB}{2Mc}L_Z \tag{30}$$

where L_Z is given by Equation (27).

$$\begin{aligned} V = & \frac{e^2B^2}{8c^2}\left(\frac{1}{m}+\frac{Z^2}{m_N}-\frac{(Z-1)^2}{M}\right)\left(X^2+Y^2\right)+\frac{(Z-1)eB}{2Mc}L_Z\frac{P^2}{2M^2c^2}-\frac{P^4}{8M^3c^2} \\ & +\left(\frac{p^2}{2\mu}-\frac{Ze^2}{q}\right)\frac{P^2}{2M^2c^2}\left(1-2\frac{m}{M}\right)-\frac{p^2}{2\mu}\frac{m}{M}\frac{P^2}{M^2c^2}-\frac{p_z^2}{2\mu}\frac{P^2}{2M^2c^2} \\ & +\frac{e^2B^2}{8mc^2}\left(q^2-q_z^2\right)\left(\frac{m_N^2}{M^2}+Z^2\frac{m}{m_N}\frac{m^2}{M^2}\right)+\frac{eB}{2mc}\left(\frac{m_N}{M}-Z\frac{m^2}{m_NM}\right)l_z+\frac{eB}{mc}\sigma_{ez} \\ & -\frac{eB}{2mc}\frac{p^2}{2m}l_z\left(\frac{m_N}{M}\frac{1}{mc^2}+\frac{m_N-m}{m_N}\frac{1}{Mc^2}\right)-\frac{ZeB}{2m_Nc}\frac{p^2}{2\mu}\left(\frac{m_N}{M}-\frac{m}{m_N}\right)l_z\frac{1}{Mc^2} \\ & -\frac{eB}{2mc}\frac{Ze^2}{q}\frac{l_z}{2Mc^2}\left(1-\frac{m_N-m}{M}-Z\frac{m}{m_N}\left(\frac{3}{2}+\frac{m}{2M}\right)\right) \\ & -\frac{eB}{2mc}\frac{p^2-p_z^2}{2m}\frac{\sigma_{ez}}{2Mc^2}\left(1+Z\frac{m}{m_N}\right)-\frac{ZeB}{2m_Nc}\frac{Ze^2}{q^3}\left(q^2-q_z^2\right)\frac{\sigma_{ez}}{Mc^2} \\ & -\frac{eB}{2mc}l_z\frac{P^2}{4M^2c^2}\left(1-3\frac{m}{M}-Z\frac{m}{m_N}\left(\frac{3}{2}+\frac{m}{2M}\right)\right) \\ & -\frac{eB}{2mc}\sigma_{ez}\frac{P^2}{2M^2c^2}\left(1+\frac{m}{M}-Z\frac{m}{M}\right) \end{aligned}$$

$$-\frac{eB}{2mc}\frac{p^2}{2m}\frac{L_Z}{Mc^2}\left(1-Z\frac{m^2}{m_N^2}\right)+\frac{eB}{2mc}\frac{Ze^2}{q}\frac{L_Z}{2Mc^2}\left(1-Z\frac{m}{m_N}\right) \tag{31}$$

There are three types of terms in the expression for V given by Equation (31):

1) terms which depend only on the CM variables, contained as the first two lines of the right hand side of Equation (31)

2) the terms which depend only on the internal variables which are the remaining terms in Equation (31), except for the third, the fourth and the last four lines and

3) coupled terms which depend on the internal as well as the c.m. variables, namely, the third, the fourth and the last four lines of Equation (31). The fact that there are coupled terms is not surprising. When the composite system with a net electric charge is in an external magnetic field, the total momentum $\boldsymbol{P}$ is not conserved and we can not go to a frame where $\boldsymbol{P}=0$ for all times and so we do not expect the Hamiltonian to be uncoupled between internal and c.m. variables, as in the case of an isolated composite system.

For future reference, we will write the perturbation V as

$$V=V_0+V_B \tag{32}$$

where V_0 is the sum of terms in V which are independent of B and V_B is the sum of terms which depend on B.

Since the unperturbed Hamiltonian H_0 is the sum of the internal Hamiltonian $h^{(0)}+h^{(1)}$ and the Landau Hamiltonian H_0^L, we will take the basis states for the first order degenerate perturbation theory calculations, the direct product of the simultaneous eigenstates of $h^{(0)}+h^{(1)}$, j^2, j_z, l^2 and s^2, and the coherent states of the Landau Hamiltonian, described above. The coherent state of the Landau Hamiltonian is the simultaneous eigenstate of the annihilation operators a_+ and a_- of the two dimensional simple harmonic oscillator [9], as defined in reference [9]. In this reference, we show that the coherent state is the best approximation to the classical state where the charged particle moves in a circle in the xy plane with the cyclotron angular frequency

$$\omega_c=\frac{(Z-1)e}{2Mc} \tag{33}$$

The coherent state can be represented by the ket vector $|\alpha_+\alpha_-\rangle$ where the complex numbers α_+ and α_- are the eigenvalues of the annihilation operators a_+ and a_-. The basis states for the perturbation calculations can then be written as $\left|\alpha_+,\alpha_-,n,j,m_j,l,s=\frac{1}{2}\right\rangle$, where the quantum number n, j, m_j and l have the usual meaning. For the coherent state to approximate [9] the classical state,

$$|\alpha_+|\gg 1 \tag{34}$$

In the above basis, the perturbation part of the Hamiltonian, V, will be diagonal. The expectation value of V in these basis stats will give the corrections to the energy levels of H_0, which are taken to be the sum of the eigenvalues of $h^{(0)}+h^{(1)}$ and the energy of the coherent states of the Landau Hamiltonian,

which for all practical purposes [9] is the kinetic energy of circular motion of the He^+ ion in the uniform magnetic field.

The eigenvalues of the internal Hamiltonian $h^{(0)}+h^{(1)}$ are themselves calculated in the first order perturbation theory, treating $h^{(1)}$ as a perturbation to $h^{(0)}$. So in our calculation we implicitly assume that the corrections to the energy levels due to the external magnetic field, the so called Zeeman splittings, are much smaller than the fine structure splittings induced by $h^{(0)}$. This assumption is justified for any external magnetic field whose strength is such that,

$$\mu_B B \ll (Z\alpha)^4 \mu c^2 \tag{35}$$

where μ_B is the Bohr magneton $\left(\frac{e\hbar}{2mc}\right)$ and α is the fine structure constant $\frac{e^2}{\hbar c}$. Numerically this condition is satisfied for any B whose magnitude is less than 10 Tesla or 10^5 gauss. For 10 Tesla, $\mu_B B \approx 5.79\times10^{-5}\text{eV}$ whereas $(Z\alpha)^4 \mu c^2 \approx 8.5\times10^{-3}\text{eV}$. It is important to emphasize that the corrections to the energy levels due to $h^{(0)}$ include the effects of nuclear motion to all orders in $\frac{m}{M}$, at least to order $\frac{v^2}{c^2}$.

Some important comments are in order about the B dependent terms in V of Equation (32), which we label as V_B. In the expression for V_B, we only included those terms whose time averages over one cycle of revolution of the circular motion represented by the coherent states, do not vanish. We calculated the time dependent expectation value of the c.m. operators, P_X, P_Y, P_XP_Y, P_X^2, P_Y^2, X, Y, X^2, Y^2 etc. by using the following expression for the time dependent coherent state,

$$\left|\Psi(t)\right\rangle = \mathrm{e}^{-\frac{i}{2}\omega_c t} D_+\left(\alpha_+ \mathrm{e}^{-i\omega_c t}\right) D_+\left(\alpha_-\right)\left|0\right\rangle_+ \otimes \left|0\right\rangle_- \tag{36}$$

where D_+ and D_- are unitary operators, as defined in our previous work [9]. The ket reactor $\left|0\right\rangle_+$ and $\left|0\right\rangle_-$ are the eigenstates of the number operators [9] $N_+ = a_+^\dagger a_+$ and $N_- = a_-^\dagger a_-$ with zero eigenvalues. We also note that

$$D_+^\dagger\left(\alpha_+ \mathrm{e}^{-i\omega_c t}\right) a_+ D_+\left(\alpha_+ \mathrm{e}^{-i\omega_c t}\right) = a_+ + \alpha_+ \mathrm{e}^{-i\omega_c t} \tag{37}$$

and its adjoint relation

$$D_+^\dagger\left(\alpha_+ \mathrm{e}^{-i\omega_c t}\right) a_+^\dagger D_+\left(\alpha_+ \mathrm{e}^{-i\omega_c t}\right) = a_+^\dagger + \alpha_+^* \mathrm{e}^{+i\omega_c t} \tag{38}$$

Using Equation (37) and Equation (38) and the expressions for X and Y in terms of a_+ and a_- we would obtain,

$$\left\langle\Psi(t)\right| X - X_{0c} \left|\Psi(t)\right\rangle = \sqrt{\frac{2\hbar}{M\omega_c}}\left|\alpha_+\right| \sin\left(\omega_c t + \delta\right) \tag{39}$$

$$\left\langle\Psi(t)\right| Y - Y_{0c} \left|\Psi(t)\right\rangle = \sqrt{\frac{2\hbar}{M\omega_c}}\left|\alpha_+\right| \cos\left(\omega_c t + \delta\right) \tag{40}$$

where X_{0c} and Y_{0c} are the x and y coordinates of the center of the circular

orbit which are given by the expressions

$$X_{0c} = \sqrt{\frac{2\hbar}{M\omega_c}}\,\mathrm{Re}(\alpha_-) \tag{41}$$

$$Y_{0c} = -\sqrt{\frac{2\hbar}{M\omega_c}}\,\mathrm{Im}(\alpha_-) \tag{42}$$

$$\sin(\delta) = \frac{\mathrm{Re}(\alpha_+)}{|\alpha_+|} \tag{43}$$

$$\cos(\delta) = \frac{\mathrm{Im}(\alpha_+)}{|\alpha_+|} \tag{44}$$

We also note that the Landau Hamiltonian H_0^L of Equation (30), in terms of a_+ and $a_+^\dagger$ becomes [9],

$$H_0^L = \hbar\omega_c\left(a_+^\dagger a_+ + \frac{1}{2}\right). \tag{45}$$

The expectation value of H_0^L in the coherent state $|\alpha_+\alpha_-\rangle$ is,

$$E_0^L = E_{CM} = \hbar\omega_c\left(|\alpha_+|^2 + \frac{1}{2}\right) \tag{46}$$

So the energy of the coherent state depends only on $|\alpha_+|$ and completely independent of α_-. So for a given expectation value of the Landau Hamiltonian, there can be an infinite number of values for α_-, corresponding to the infinite degeneracy of the Landau energy levels. Equations (41)-(44) suggest that this infinite degeneracy of the coherent state for a given $|\alpha_+|$, correspond to an infinite number of possible centers of the circle and the different phases of α_+ representing different phases of the circular motion, for a given radius and a given energy of the circular motion. We have the freedom to choose he center of the circular orbit at the origin of the xy coordinate system and from Equation (41) and Equation (42) if follows that

$$\alpha_- = 0 \tag{47}$$

with this choice we will find,

$$\langle\Psi(t)|X|\Psi(t)\rangle = \sqrt{\frac{2\hbar}{M\omega_c}}|\alpha_+|\sin(\omega_c t + \delta) \tag{48}$$

$$\langle\Psi(t)|Y|\Psi(t)\rangle = \sqrt{\frac{2\hbar}{M\omega_c}}|\alpha_+|\cos(\omega_c t + \delta) \tag{49}$$

$$\langle\Psi(t)|X^2|\Psi(t)\rangle = \frac{\hbar}{M\omega_c}\left[1 + |\alpha_+|^2\left(1 + \cos(2\omega_c t)\right)\right] \tag{50}$$

$$\langle\Psi(t)|Y^2|\Psi(t)\rangle = \frac{\hbar}{M\omega_c}\left[1 + |\alpha_+|^2\left(1 - \cos(2\omega_c t)\right)\right] \tag{51}$$

$$\langle\Psi(t)|X^2 + Y^2|\Psi(t)\rangle = \frac{2\hbar}{M\omega_c}\left(1 + |\alpha_+|^2\right) \tag{52}$$

$$\left\langle \Psi(t)\left| P_X \right| \Psi(t)\right\rangle = \sqrt{\frac{M\omega_c\hbar}{2}}\left|\alpha_+\right|\cos\left(\omega_c t+\delta\right) \tag{53}$$

$$\left\langle \Psi(t)\left| P_Y \right| \Psi(t)\right\rangle = -\sqrt{\frac{M\omega_c\hbar}{2}}\left|\alpha_+\right|\sin\left(\omega_c t+\delta\right) \tag{54}$$

$$\left\langle \Psi(t)\left| P_X^2 \right| \Psi(t)\right\rangle = \frac{M\omega_c\hbar}{4}\left[1+\left|\alpha_+\right|^2\left(1-\cos\left(2\omega_c t\right)\right)\right] \tag{55}$$

$$\left\langle \Psi(t)\left| P_Y^2 \right| \Psi(t)\right\rangle = \frac{M\omega_c\hbar}{4}\left[1+\left|\alpha_+\right|^2\left(1+\cos\left(2\omega_c t\right)\right)\right] \tag{56}$$

$$\left\langle \Psi(t)\left| P_X^2+P_Y^2 \right| \Psi(t)\right\rangle = \frac{M\omega_c\hbar}{2}\left(1+\left|\alpha_+\right|^2\right) \tag{57}$$

$$\left\langle \Psi(t)\left| P_X P_Y \right| \Psi(t)\right\rangle = \frac{M\omega_c\hbar}{4}\left[\mathrm{Re}\left(\alpha_+\right)^2\sin\left(2\omega_c t\right)-\mathrm{Im}\left(\alpha_+\right)^2\cos\left(2\omega_c t\right)\right] \tag{58}$$

From Equation (48), Equation (49), Equation (53), Equation (54) and Equation (58) it is clear that the time averages of P_X, P_Y, X, Y and $P_X P_Y$ will vanish over any cycle of circular revolution. Since the cyclotron frequency of revolution $\omega_c = \dfrac{(Z-1)eB}{Mc}$ is of the order of 10^{10} s^{-1} even for B as small as 1000 gauss, what we can observe for the CM variables of the circular motion will be time averages. This is why we neglected terms in the perturbation Hamiltonian V_B whose time average over one cycle of revolution vanished.

4. Results of the Perturbative Calculations in the First Order

We now give the results of the first order perturbative results where V of equation (31) is the perturbation and the basis states are the direct product of the eigenstates of $h^{(0)}+h^{(1)}$, of the Equation (28) and Equation (29) and the coherent states of the Landau Hamiltonian of Equation (30). We can write the first order results which depend only on the quantum numbers n, l, j and m_j as well as the magnetic field B and E_{CM}, as,

$$E_{nljm_j} = E_{nlj}^{(0)} + \Delta E_{nljm_j}^{(1)}\left(B\right) \tag{59}$$

where $E_{nlj}^{(0)}$ are the eigenvalues of the internal Hamiltonian $h^{(0)}+h^{(1)}$ to first order in $h^{(1)}$, and $\Delta E_{nljm_j}^{(1)}\left(B\right)$ are the first order corrections due to the perturbation V of Equation (31) which depend on the external uniform magnetic field B.

$$E_{nlj}^{(0)} = -Z^2\alpha^2\frac{1}{2}\mu c^2\frac{1}{n^2}-\frac{1}{2}\left(Z\alpha\right)^4\mu c^2\frac{1}{n^3}\left(\frac{m_N}{M}\right)^2\left(\frac{1}{j+\frac{1}{2}}-\frac{3}{4n}\right)$$

$$+\frac{1}{2}\left(Z\alpha\right)^4\mu c^2\frac{1}{n^3}\left(\frac{m_N}{M}\right)^2\frac{m}{M}\left(\frac{1}{j}-\frac{3}{4n}\right)$$

$$-\frac{1}{2}\left(Z\alpha\right)^4\mu c^2\left(\frac{m}{M}\right)^3\left[\frac{1}{n^3\left(l+\frac{1}{2}\right)}-\frac{3}{4n^4}\right]$$

$$+\frac{1}{2}(Z\alpha)^4\,\mu c^2\,\frac{1}{n^3}\frac{mm_N}{M^2}\frac{1}{l\left(l+\frac{1}{2}\right)(l+1)}$$

$$\times\left[j(j+1)-\frac{3}{4}-l(l+1)\right](1-\delta_{l0})+\frac{1}{2}(Z\alpha)^4\,\mu c^2\,\frac{1}{n^4}\frac{mm_N}{M^2}\left(1-\frac{2n}{l+\frac{1}{2}}\right) \quad (60)$$

$$-\frac{Ze^2}{2mm_Nc^2}\hbar^2\left[2\int_0^\infty R_{nl}(r)\frac{\mathrm{d}R_{nl}}{\mathrm{d}r}\mathrm{d}r+\int_0^\infty R_{nl}(r)\frac{\mathrm{d}^2R_{nl}}{\mathrm{d}^2r}r\mathrm{d}r\right]+E_{CM}$$

$$\Delta E^{(1)}_{nljm_j}(B,E_{CM})$$

$$=-\frac{1}{4}Z^2\alpha^2\,\frac{\mu c^2}{n^2}\left(\frac{E_{CM}}{M^2c^2}\right)\left(\frac{3}{2}-Z+\frac{l(l+1)-m_j^2-\frac{3}{4}}{(2l-1)(2l+3)}+\frac{2m_j^2}{(2l-1)(2l+1)(2l+3)}\right)$$

$$+\frac{e^2B^2}{8M^2c^2}\left(\frac{m_N^2}{m}+Z^2\frac{m^2}{m_N}\right)\frac{a_0^2n^2}{2Z^2}\left[5n^2+1-3l(l+1)\right]$$

$$\times\left(1-2\frac{l(l+1)-m_j^2-\frac{3}{4}}{(2l-1)(2l+3)}-\frac{4m_j^2}{(2l-1)(2l+1)(2l+3)}\right)$$

$$+\mu_BB\left(1-\frac{m}{M}\right)\left(1-Z\frac{m^2}{m_N^2}\right)\times\frac{j(j+1)+l(l+1)-\frac{3}{4}}{2j(j+1)}m_j$$

$$+\mu_BB\frac{j(j+1)+\frac{3}{4}-l(l+1)}{j(j+1)}m_j-\mu_BBZ^2\alpha^2\frac{1}{n^2}\frac{j(j+1)+l(l+1)-\frac{3}{4}}{4j(j+1)}$$

$$\times\left[\frac{m_N^2}{M^2}\left(1-\frac{m^2}{m_NM}\right)+Z\frac{m^2}{M^2}\left(\frac{m_N}{M}-\frac{m}{m_N}\right)\right]$$

$$-\mu_BBZ^2\alpha^2\frac{1}{n^2}\left[2\frac{m}{M}-\frac{m}{m_N}\left(\frac{3}{2}+\frac{m}{2M}\right)Z\right]\frac{m_N}{M}\frac{m}{M}\frac{j(j+1)+l(l+1)-\frac{3}{4}}{4j(j+1)}m_j$$

$$-\mu_BBZ^2\alpha^2\frac{1}{n^2}\frac{1}{4}\left(\frac{m_N}{M}\right)^2\frac{m}{M}\left(1+Z\frac{m}{m_N}\right)$$

$$\times\left(\frac{j(j+1)+\frac{3}{4}-l(l+1)}{2j(j+1)}-\frac{2\left(j^2+j-m_j^2-1\right)}{(2l-1)(2l+1)(2l+3)}\right)m_j$$

$$-Z\mu_BBZ^2\alpha^2\,\frac{1}{n^2}\left(\frac{m}{M}\right)^2\left\{1-2\left[l(l+1)-m_j^2-\frac{3}{4}\right]\right.$$

$$\left.-\frac{4m_j^2}{(2l-1)(2l+1)(2l+3)}\right\}\times\frac{j(j+1)+\frac{3}{4}-l(l+1)}{2j(j+1)}m_j$$

$$-\mu_BB\left(\frac{E_{CM}}{Mc^2}\right)\left[1-3\frac{m}{M}-Z\frac{m}{m_N}\left(\frac{3}{2}+\frac{m}{2M}\right)\right]$$

$$\times\frac{j(j+1)+l(l+1)-\frac{3}{4}}{4j(j+1)}m_j-\mu_B B\left(\frac{E_{CM}}{Mc^2}\right)\left(1+\frac{m}{M}-Z\frac{m}{M}\right)$$
$$\times\frac{j(j+1)+\frac{3}{4}-l(l+1)}{4j(j+1)}m_j-\frac{(Z-1)}{2}\left(\frac{E_{CM}}{Mc^2}\right)E_{CM}-\frac{5}{8}\left(\frac{E_{CM}}{Mc^2}\right)E_{CM} \quad (61)$$

Several comments are in order about Equation (60) and Equation (61). In these equations E_{CM} is the energy of the circular motion of the C.M. of the He^+ ion, corresponding to the coherent state of the Landau Hamiltonian of Equation (30). Its explicit expression is given by Equation (46). Since the CM motion represented by the coherent state [9] is practically classical, the eigenvalue of the annihilation operator a_+ satisfies the condition,

$$|\alpha_+|\gg 1 \quad (62)$$

which means the $\frac{1}{2}$ in Equation (46) can be neglected. So,

$$|\alpha_+|^2\approx\frac{E_{CM}}{\hbar\omega_c} \quad (63)$$

where ω_c is given by Equation (33). In obtaining Equation (61) we also made use of the relation,

$$\begin{aligned}\langle\alpha_+\alpha_-|L_Z|\alpha_+\alpha_-\rangle&=\langle\alpha_+\alpha_-|\hbar\left(a_-^\dagger a_- - a_+^\dagger a_+\right)|\alpha_+\alpha_-\rangle\\&=\hbar\left(|\alpha_-|^2-|\alpha_+|^2\right)=-\hbar|\alpha_+|^2\end{aligned} \quad (64)$$

In obtaining the last equality in Equation (64) we have put $\alpha_-=0$, by choosing the center of the circular orbit at the origin of the xy plane. We also made use of the relation,

$$\langle\alpha_+\alpha_-|L_Z\frac{P^2}{4M^2c^2}|\alpha_+\alpha_-\rangle=\frac{M\hbar\omega_c}{2}\frac{|\alpha_+|^4}{4M^2c^2}=\frac{\hbar\omega_c}{8}\frac{|\alpha_+|^4}{Mc^2}=\frac{E_{CM}^2}{8\hbar\omega_c Mc^2} \quad (65)$$

Also, in Equation (61) the symbol μ_B represents the Bohr magneton of the electron. In regular Gaussian units, where $\hbar\neq 1$

$$\mu_B=\frac{e\hbar}{2mc}=5.79\times10^{-5}\text{eV}\cdot\text{T}^{-1}=5.79\times10^{-9}\text{eV}\cdot\text{G}^{-1} \quad (66)$$

In deriving Equation (31) and Equation (61) we included terms quadratic in B, only in the nonrelativistic terms, but not in the first order relativistic correction, for reasons mentioned before. There is only one term quadratic in B in the expression for $E^{(1)}_{nljm_j}$ of Equation (61). This term is of order $\left(\frac{\mu_B B}{\alpha^2\mu c^2}\right)$ compared to the dominant term for Zeeman splitting, which are the first two terms proportional to $\mu_B B$ in Equation (61). For a magnetic field B even of the order of 10 Tesla, $\left(\frac{\mu_B B}{\alpha^2\mu c^2}\right)$ is of the order of 10^{-2}, which is really small compared to the leading term.

The results of Equation (61) includes the usual nonrelativistic result for the

Zeeman effect, including the effects of nuclear motion, plus all relativistic corrections of order $\frac{v^2}{c^2}$. The first two terms proportional to $\mu_B B$ in Equation (61) goes over to the usual nonrelativistic results in text books [16] if we let $\frac{m_N}{M} \to 1$ or equivalently $\frac{m}{M} \to 0$. The dominant relativistic correction is the first term proportional $(\mu_B B) Z^2 \alpha^2$ in Equation (61). The other terms proportional to $(\mu_B B) Z^2 \alpha^2$ are reduced by factors of $\left(\frac{m}{M}\right)$ or higher powers of $\left(\frac{m}{M}\right)$. What is really new in our approach is the correction to the Zeeman splitting due to the circular motion of the center of mass of the He^+ ion given by the terms proportional to $(\mu_B B)\left(\frac{E_{CM}}{Mc^2}\right)$ in Equation (61). These corrections are really small, even smaller than the dominant $Z^2\alpha^2$ correction term, for nonrelativistic energies of the circular motion, which we have assumed in this paper. Nevertheless, they are interesting and can be checked in accurate experiments. Of course these terms will vanish if $E_{CM} = 0$ or if the velocity of the He^+ ion is directed entirely along the direction of the applied magnetic field.

The last term in Equation (60) and the last two terms in Equation (61) which depend only on E_{CM} is like a constant addition to the energies and does not affect the transition energies and hence frequencies of spectral lines and so may be neglected.

We should also make some important comments about Equation (60) which gives the fine structure splittings of the He^+ ion, including the effects of nuclear motion. If we let $\frac{m}{M} \to 0$ or equivalently $\frac{m_N}{M} \to 1$, our results will go over into the usual text book results [16]. In Equation (60), $R_{nl}(r)$ is the radial part of the eigenfunction of $h^{(0)}$, which are the well-known product of the exponential function and the associated Laguerre polynomials [16]. It can be shown,

$$\int_0^\infty R_{nl}(r)\frac{\mathrm{d}R_{nl}}{\mathrm{d}r}\mathrm{d}r = -\frac{4Z^3}{a_0^3}\frac{1}{n^4}\frac{1}{\left[(n+l)!\right]^2} \tag{67}$$

The other integral involving the second derivative of $R_{nl}(r)$ is too long to write down. But from dimensional arguments, it is also proportional to $\frac{1}{a_0^3}$ where a_0 is the Bohr radius. So the contribution from both integral terms in Equation (60) to the energy levels is of the order of $\frac{1}{Z^3}\left(\frac{m}{M}\right)$ times the dominant fine structure term, which is the first term proportional to $(Z\alpha)^4$ in Equation (60), and hence utterly negligible, compared to it. The integral terms in Equation (60) come from the magnetic interaction between the moving electron and the magnetic field produced by the moving He^4 nucleus, and it will of course vanish in the limit of the infinite mass for the nucleus.

Another important point we should make about our calculations is that we

have chosen the Z-component of the CM momentum $\boldsymbol{P}$, namely,

$$P_Z = 0 \tag{68}$$

In the general case, when $P_Z \neq 0$, the He^+ ion is moving with a uniform velocity along the Z-axis, since there is no external force in the Z-direction. So the results in the general case when $P_Z \neq 0$, can be obtained by a Lorentz transformation to the appropriate order in $\frac{v}{c}$, of the total energy, including the rest mass energy. So the results for $P_Z \neq 0$ can be obtained from our expressions in Equation (60) and Equation (61) by adding the term, $\frac{P_Z^2}{2M}$ to E_{CM} on the right hand side of Equation (60) and multiplying the remaining terms in Equation (60) by $\left(1+\frac{P_Z^2}{2M^2c^2}\right)$. We should also add $\frac{P_Z^4}{8M^3c^2}$ to the right hand side of Equation (61) while multiplying all the term on the right hand side of Equation (61), except the last two terms involving only E_{CM}, by the factor $\left(1+\frac{P_Z^2}{2M^2c^2}\right)$.

5. Summary and Concluding Remarks

We have calculated the order $\frac{v^2}{c^2}$ relativistic correction to the Zeeman Effect in He^+ ion, when its CM moves in a circular orbit under the action of a uniform magnetic field perpendicular to its plane of motion. We have assumed the weak field approximation in the sense that the splittings of energy levels due to the magnetic field are much smaller than the fine structure splittings. This is justified so long as the strength of the magnetic field is less than two Tesla or 20,000 Gauss. We also assumed the CM circular motion is approximately nonrelativistic. Even if the kinetic energy of the CM circular motion is 1 MeV, it is much less than the rest mass energy of the He^+ ion which is about $3.76\times10^9\text{eV}$.

What is novel in our approach is that we have chosen the basis states, for the first order degenerate perturbation theory, to be the direct product of the coherent states $|\alpha_+\alpha_-\rangle$ of the Landau Hamiltonian H_0^L of Equation (30) and the internal state which is a simultaneous eigenstate of the internal nonrelativistic Hamiltonian $h^{(0)}$, j^2, l^2, s^2 and j_Z, namely $|nljm_j\rangle$. The coherent state is an excellent approximation [9] to the classical state of circular motion, so long as the magnitude of the eigenvalue of the annihilation operator a_+ of the Landau Hamiltonian H_0^L is much greater than one. For example, for $E_{CM} = 10\,\text{keV}$, $|\alpha_+| \approx 10^4$, when $E_{CM} = 1\,\text{MeV}$, $|\alpha_+| \approx 10^5$, and they suggest excellent classical approximations. Since the He^+ ion has a net electric charge the total momentum is not conserved and so we do not expect complete separation of CM and internal motion. In fact the coupling terms between CM and internal motion give corrections of order $\frac{E_{CM}}{Mc^2}$ to the Bohr energy levels which should be detectable in careful spectroscopic experiments involving He^+ ion in a uniform magnetic

field, even though $\frac{E_{CM}}{Mc^2}$ is of the order of 10^{-5}, when E_{CM} is of the order of 10 keV.

The study of Zeeman effect in He^+ ion is important in astrophysical situation [17] [18], where there are He^+ ions in magnetic fields. We have only considered He^4 ion, where the He^4 nucleus has zero spin and zero magnetic moment. But our treatment can be easily extended to He^3 ion, where He^3 nucleus has spin $\frac{1}{2}$ and a spin magnetic moment. In this case there will be extra terms in the Hamiltonian of Equation (1) and hence in Equation (22) which will give extra terms in the expressions of Equation (60) and Equation (61). Equation (60) will then include hyperfine splittings in the energy levels, due to the interaction of the electronic and the nuclear spin magnetic moments. The study of composite systems with net electric charge in a magnetic field has also received considerable attention [19] [20] in atomic and solid state physics.

Conflicts of Interest

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

References

[1] Baye, D. (1982) *Journal of Physics B*: *Atomic and Molecular Physics*, **15**, L795. https://doi.org/10.1088/0022-3700/15/22/003

[2] Baye, D. and Vincke, M. (1986) *Journal of Physics B*: *Atomic and Molecular Physics*, **19**, 4051. https://doi.org/10.1088/0022-3700/19/24/009

[3] Avron, I.W. and Simon, B. (1978) *Annals of Physics*, **114**, 431. https://doi.org/10.1016/0003-4916(78)90276-2

[4] Baye, D. (1983) *Journal of Physics A*: *Mathematical and General*, **16**, 3207. https://doi.org/10.1088/0305-4470/16/14/016

[5] Dickinson, A.S. and Patterson, J.M. (1986) *Journal of Physics A*: *Mathematical and General*, **19**, 1811. https://doi.org/10.1088/0305-4470/19/10/022

[6] Herold, H.R.H. and Wunner, G. (1981) *Journal of Physics B*: *Atomic and Molecular Physics*, **14**, 751. https://doi.org/10.1088/0022-3700/14/4/022

[7] Thomas, I.L. (1971) *Physical Review A*, **3**, 1022. https://doi.org/10.1103/PhysRevA.3.1022

[8] Johnson, J.O. and Yang, K.-H. (1983) *Reviews of Modern Physics*, **55**, 109. https://doi.org/10.1103/RevModPhys.55.109

[9] Li, W. and Sebastian, K. (2018) *European Journal of Physics*, **39**, Article ID: 045403. https://doi.org/10.1088/1361-6404/aab985

[10] Krajcik, R.A. and Foldy, L.L. (1974) *Physical Review D*, **10**, 1777. https://doi.org/10.1103/PhysRevD.10.1777

[11] Sebastian, K.J. and Yum, D. (1979) *Physical Review D*, **19**, 2509. https://doi.org/10.1103/PhysRevD.19.2509

[12] Close, F.E. and Osborn, H. (1970) *Physical Review D*, **2**, 2127. https://doi.org/10.1103/PhysRevD.2.2127

[13] Close, F.E. and Copley, L.A. (1970) *Physical Review B*, **19**, 477. https://doi.org/10.1016/0550-3213(70)90362-7

[14] Sebastian, K.J. (1981) *Physical Review A*, **23**, 2810. https://doi.org/10.1103/PhysRevA.23.2810

[15] Krajcik, R.A. and Foldy, L.L. (1970) *Physical Review Letters*, **24**, 545. https://doi.org/10.1103/PhysRevLett.24.545

[16] Shankar, R. (1994) Principles of Quantum Mechanics. 2nd Edition, Springer, Berlin. https://doi.org/10.1007/978-1-4757-0576-8

[17] Richard, B.L., Crutcher, M., Troland, T.H. and Kazes, I. (1999) *The Astrophysical Journal*, **514**, L121. https://doi.org/10.1086/311952

[18] Cooksy, G.A. and Saykally, R.J. (1986) *The Astrophysical Journal*, **305**, L89. https://doi.org/10.1086/184691

[19] Euna Jo, J.H., An, K. and Lee, S. (2011) *Physical Review B*, **84**, Article ID: 174423.

[20] van der Laan, G. and Thole, B.T. (1991) *Physical Review B*, **43**, 13401. https://doi.org/10.1103/PhysRevB.43.13401

Journal of Modern Physics, 2020, 11, 475-485
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Dualism of the Heisenberg and Schrödinger Approaches to the Quantum States Entering a One-Dimensional Electron Gas

Stanisław Olszewski

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

How to cite this paper: Olszewski, S. (2020) Dualism of the Heisenberg and Schrödinger Approaches to the Quantum States Entering a One-Dimensional Electron Gas. *Journal of Modern Physics*, **11**, 475-485.
https://doi.org/10.4236/jmp.2020.113030

Received: February 19, 2020
Accepted: March 22, 2020
Published: March 25, 2020

Abstract

According to quantum mechanics, the commutation property of the energy Hamiltonian with the momentum operator should give the definite values not only for energy but also for the momentum quantum levels. A difficulty provided by the standing-like boundary conditions of the electron gas is that the Hamiltonian eigenfunctions are different than eigenfunctions of the momentum operator. In results the electron momenta are obtained from the correspondence rule between the classical and quantum mechanics given by Landau and Lifshits. As a consequence the statistics of solutions representing not only the energy values but also the electron momenta should be taken into account. In the Heisenberg picture of quantum mechanics, the momenta are easily obtained because the electron oscillators are there directly considered. In fact, the Hamiltonian entering the Heisenberg method can be defined in two different ways each giving the set of the electron energies known from the Schrödinger's approach.

Keywords

Fundamentals of the Modern Quantum Theory, Heisenberg Picture, Its Momentum Results and the Energy Matrix, Schrödinger Picture and Its Energy Results

1. Introduction

Historically we had a competition of the Schrödinger and Heisenberg formalisms in their approach to develop the modern quantum mechanics. Certainly, the Schrödinger method [1] [2] [3] [4] occurred to be more practical in calculating the physical properties of numerous electron systems. But, on the other hand, Heisenberg was the first one who presented foundations of the idea of the

observables and operators belonging to them [5] [6].

Nevertheless, a practical background of the Heisenberg theory remained rather poor. In fact only the oscillators of different kinds and their properties, especially those concerning the behaviour of the matrix elements, are dominating in the Heisenberg approach.

Our idea is to compare the Heisenberg and Schrödinger methods in examining a very simple system represented by the electrons enclosed in a one-dimensional potential box. To the author's knowledge this comparison seems to have never been done before. A wave-mechanical treatment of such a system performed according to the Schrödinger method occurred to be very simple. On the other hand, the oscillating character and properties of the system remained fully neglected.

An alternative method, suitable for the oscillatory examination, became the Heisenberg theory and an approach to it presented in [7]. The both methods of quantum mechanics could be examined for a non-interacting electron gas case. Physically they indicate a different behaviour of electrons in each kind of the examined theory, as well as in course of their application in statistics.

We note that the Schrödinger's approach to the one-dimensional electron gas can be done in a unique way. On the other hand, the Heisenberg theory could be applied to the same gas in two different ways, each basing on a different Hamiltonian formula used for the same gas object. In consequence the statistics leading to the states occupation by the electrons in the Schrödinger theory should be also slightly changed respectively to the examined Heisenberg's Hamiltonian case.

This is so because of the boundary conditions imposed on electrons entering the Schrödinger model. In effect the electron wave functions represent the standing waves and they are not the eigenfunctions of the momentum operator, contrary to situation due to the Bloch's boundary conditions applied usually to solids. In this second case, however, the requirement of a finite and strictly one-dimensional gas model cannot be satisfied.

A difficulty with the boundary conditions entering the Schrödinger method can be removed when the correspondence rule due to Landau and Lifshits [8] concerning the classical and quantum approaches to the mechanics of electron motion is applied. Because the free-electron Hamiltonian commutes with the operator of the electron momentum, the stationary states of energy should be accompanied by the stationary states of momentum. This is easily obtained on a semi-classical way when the Landau-Lifshits rule is assumed to hold. In effect we should obtain the electron momenta corresponding to the quantum energy levels given by the Schrödinger method, on condition the quantum number n of energy is large. In fact the electron momenta obtained with the aid of the Heisenberg method are found identical with those deduced with the aid of the Schrödinger model also for the low quantum numbers n; see Sec. 5. This imposes a question of statistics with which both the energy and momenta of free-electron particles can be satisfactorily considered.

In brief, an alternative method to Schrödinger's—much suitable for the oscillatory examination—is that of Heisenberg; see e.g. [7] [9].

2. Schrödinger's Approach to the Free-Electron Particles

This approach represents an elementary Schrödinger quantization process and its parameters; see e.g. [10].

Within a one-dimensional potential box of length L we have the free electrons whose energies are defined by the eigenvalues of the Hamiltonian operator:

$$\hat{H} = -\frac{\hbar^2}{2m_e}\frac{\mathrm{d}^2}{\mathrm{d}x^2} = \frac{\hat{p}_x^2}{2m_e}; \tag{1}$$

$\hat{p}_x$ is the electron momentum operator, m_e is the electron mass. In effect we have the eigenequation:

$$\hat{H}\psi_n(x) = E_n\psi_n(x). \tag{2}$$

The wave functions $\psi_n(x)$ satisfy the following boundary conditions at the box ends, viz.

$$\psi_n(x=0) = \psi_n(x=L) = 0. \tag{3}$$

Suitable $\psi_n(x)$ are easily verified to be

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}x\right), \tag{4}$$

the eigenenergies are (see e.g. [10]):

$$E_n = \frac{n^2h^2}{8m_eL^2}; \tag{5}$$

Because of (3) the wave functions (4) are called the standing-like wave functions in the potential box.

3. Oscillating Character of Electrons Described by ψ_n and E_n

Free electrons should satisfy the equation

$$E_n = \frac{m_e}{2}v_n^2 \tag{6}$$

where the absolute value of the electron velocity v_n is

$$v_n = \left(\frac{2E_n}{m_e}\right)^{1/2} = \left(\frac{n^2h^2}{4m_e^2L^2}\right)^{1/2} = \frac{nh}{2m_eL}. \tag{7}$$

This value of the velocity is expected to be dominating in course of the electron oscillation within the interval

$$0 < x < L. \tag{8}$$

Evidently the electron motion is going from $x=0$ to $x=L$ and vice versa.

The time period T_n of the oscillation satisfies the formula

$$\frac{2L}{T_n} = v_n, \tag{9}$$

so

$$T_n = \frac{2L}{v_n} = 2L \times \frac{2m_e L}{nh} = \frac{4m_e L^2}{nh}. \tag{10}$$

The motion frequency ν due to T_n is

$$\nu_n = \frac{1}{T_n} = \frac{nh}{4m_e L^2}. \tag{11}$$

In effect the energy provided by ν_n becomes

$$E_n^{osc} = h\nu_n = \frac{nh^2}{4m_e L^2} = \hbar\omega_n \tag{12}$$

so

$$\omega_n = \frac{nh^2}{4m_e L^2}\frac{1}{\hbar} = \frac{\pi nh}{2m_e L^2} \tag{13}$$

and its reversal is

$$\frac{1}{\omega_n} = \frac{2m_e L^2}{\pi nh} \tag{14}$$

Now our idea is to match results of Sec. 2 with those of Sec. 3.

Before we do such comparison let us note that parameters T_n and v_n satisfy the original Heisenberg relation [5]:

$$\frac{d}{dn}(nh) = h = \frac{d}{dn}\oint m_e \dot{x}^2 dt \tag{14a}$$

For, from (10) and because of $\dot{x} = v_n$ given in (9), we obtain for (14a):

$$\frac{d}{dn}\oint m_e \dot{x}^2 dt = \frac{d}{dn}\left(m_e v_n^2 T_n\right) = \frac{d}{dn}\left(m_e \frac{4L^2}{T_n}\right) = \frac{d}{dn}\left(m_e \frac{nh}{m_e}\right) = h \tag{14b}$$

which is identical to the result in [8] and (14a).

In the same way we have

$$\oint pdq = \oint m_e \dot{x} dx = m_e v_n 2L = m_e \frac{(2L)^2}{T_n} = m_e 4L^2 \frac{nh}{4m_e L^2} = nh. \tag{14c}$$

If the result in (14c) is considered as the action J, it becomes evident that the derivative of energy in (5) done with respect to J provides us with the electron oscillation frequency (11).

4. Heisenberg Approach Applied to the Electron Oscillators

In the first step of the Heisenberg approach to the electron gas enclosed in a potential box we consider the Hamiltonian of an oscillator moving in direction of the axis x [7]:

$$\hat{H} = \frac{\hat{p}_x^2}{2m_e} + \frac{m_e\omega^2}{2}x^2 = m_e\omega^2 \begin{Vmatrix} x_{01}x_{10} & 0 & 0 & \cdots \\ 0 & x_{01}x_{10} + x_{12}x_{21} & 0 & \cdots \\ 0 & 0 & x_{12}x_{21} + x_{23}x_{32} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{Vmatrix} \tag{15}$$

where

$$x_{01}x_{10} = \frac{\hbar}{2m_e\omega}, \tag{16}$$

$$x_{12}x_{21} = 2\frac{\hbar}{2m_e\omega}, \tag{16a}$$

$$x_{23}x_{32} = 3\frac{\hbar}{2m_e\omega}, \tag{16b}$$

$$x_{34}x_{43} = 4\frac{\hbar}{2m_e\omega}, \tag{16c}$$

$$\vdots$$

The x_{mn} are the matrix elements of x calculated between the oscillator states m and n, the frequencies ω are those calculated in (13) taken for $n=1$.

When the results in the (16) formulae are substituted to (15) they give the following diagonal elements for the matrix presented in (15):

$$x_{01}x_{10} = \frac{L^2}{2\pi^2}, \tag{17}$$

$$x_{01}x_{10} + x_{12}x_{21} = (1+2)\frac{L^2}{2\pi^2} = \frac{3L^2}{2\pi^2}, \tag{17a}$$

$$x_{12}x_{21} + x_{23}x_{32} = (2+3)\frac{L^2}{2\pi^2} = \frac{5L^2}{2\pi^2}, \tag{17b}$$

$$x_{23}x_{32} + x_{34}x_{43} = (3+4)\frac{L^2}{2\pi^2} = \frac{7L^2}{2\pi^2}, \tag{17c}$$

$$\vdots$$

which give the following result for (15):

$$\hat{H} = m_e\omega^2 \begin{Vmatrix} \frac{L^2}{2\pi^2} & 0 & 0 & 0 & \cdots \\ 0 & \frac{3L^2}{2\pi^2} & 0 & 0 & \cdots \\ 0 & 0 & \frac{5L^2}{2\pi^2} & 0 & \cdots \\ 0 & 0 & 0 & \frac{7L^2}{2\pi^2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{Vmatrix} \tag{18}$$

Because of $\omega = \omega_1$ calculated in (13), the diagonal matrix elements in (18) give the oscillator energies

$$E_m^{osc} = m_e\omega^2\frac{L^2}{2\pi^2}(2m+1) = m_e\left(\frac{\pi h}{2m_eL^2}\right)^2\frac{L^2}{2\pi^2}(2m+1) = \frac{h^2}{8m_eL^2}(2m+1) \tag{19}$$

where

$$m = 0,1,2,3,\cdots \tag{19a}$$

Let us note that the oscillator frequency ω in [7] is considered as a known parameter. In our calculations this frequency is deduced from the electron motion in the potential box; see Sec. 3.

This feature enables us to present the matrix elements entering the oscillator energy in terms of the matrix elements dependent on the properties characteris

tic for the motion in the potential box. Typically for the Heisenberg's treatment of an oscillator we choose only a single oscillation frequency ω for calculating all quantum states. In the Schrödinger picture this frequency is associated with the lowest quantum state $n=1$; see (13).

A passage to the particle energy in the box is very simple. We note that the sums entering the partial traces of the diagonal matrix elements given in (19), viz.

$$\sum_{m=0}^{m=n-1}(2m+1), \tag{20}$$

give respectively

$$\begin{aligned} n^2 &= 1^2 && \text{for } n=1,\\ n^2 &= 1+3=2^2 && \text{for } n=2,\\ n^2 &= 1+3+5=3^2 && \text{for } n=3,\\ n^2 &= 1+3+5+7=4^2 && \text{for } n=4,\\ n^2 &= 1+3+5+7+9=5^2 && \text{for } n=5,\\ &\vdots \end{aligned} \tag{21}$$

In effect the (21)—combined with the factor entering the last term in (19)—provide us with the Schrödinger energy results presented in (5).

5. An alternative Heisenberg Treatment of Free Electrons in the Potential Box

In this case we apply the Hamiltonian different than in the oscillator case examined in Sec. 4. This kind of treatment takes into account only the kinetic part of the Hamiltonian and neglects the whole of the x-dependent (potential) part in the first row of (15):

$$\hat{H}=\frac{\hat{p}_x^2}{2m_e}. \tag{22}$$

Certainly the constant parameters become different than those applied in the Hamiltonian case of Sec. 4.

The first of the diagonal terms of p_x^2 belonging to the energy matrix representing the Hamiltonian in (22) is:

$$p_{01}p_{10}=x_{01}x_{10}2m_e^2\omega^2=\frac{\hbar}{2m_e\omega}2m_e^2\omega^2=\frac{\hbar}{2}2m_e\omega=\frac{h}{2\pi}m_e\frac{\pi h}{2m_eL^2}=\frac{h^2}{4L^2}. \tag{23}$$

In fact

$$p_{01}=m_ev_1=m_e\frac{h}{2m_eL}=\frac{h}{2L} \tag{24}$$

if we note that v_1 entering (24) and

$$E_1=\frac{1}{2m_e}p_{01}p_{10}=\frac{h^2}{8m_eL^2} \tag{25}$$

calculated from (23) are equal respectively to v_1 and E_1 obtained in the

Schrödinger theory; see Sec. 3.

The next diagonal element of the matrix $\hat{p}_x^2$ is

$$p_{02}p_{20} = x_{12}x_{21}2m_e^2\omega^2 = \left(\frac{\hbar}{2m_e\omega} + \frac{2\hbar}{2m_e\omega}\right)2m_e^2\omega^2$$
$$= 3\hbar 2m_e\omega = 3\frac{h}{2\pi}m_e\frac{\pi h}{2m_eL^2} = \frac{3}{4}\frac{h^2}{L^2}. \tag{26}$$

When divided by $2m_e$ the expression (26) becomes

$$\frac{1}{2m_e}p_{02}p_{20} = \frac{3h^2}{8m_eL^2}. \tag{27}$$

This result added to that obtained in (25) gives

$$\frac{1}{2m_e}p_{01}p_{10} + \frac{1}{2m_e}p_{02}p_{20} = \frac{4h^2}{8m_eL^2} = \frac{2^2h^2}{8m_eL^2} \tag{28}$$

which is precisely the next Schrödinger value for the free-electron energy, *i.e.* it is corresponding to the quantum level $n = 2$.

If we take the next diagonal term for the matrix of $\hat{p}_x^2$ which is

$$p_{03}p_{30} = \left(\frac{\hbar}{m_e\omega} + \frac{3}{2}\frac{\hbar}{m_e\omega}\right)2\left(m_e\omega\right)^2 = \left(1+\frac{3}{2}\right)\hbar 2m_e\omega$$
$$= \frac{5\hbar}{2}2m_e\omega = 5\frac{h}{2\pi}m_e\frac{\pi h}{2m_eL^2} = \frac{5}{4}\frac{h^2}{L^2}, \tag{29}$$

it gives, when multiplied by $\frac{1}{2m_e}$, the result

$$\frac{1}{2m_e}p_{03}p_{30} = \frac{5}{8}\frac{h^2}{m_eL^2}. \tag{30}$$

The sum of terms (25), (27) and (30) becomes

$$\frac{1}{2m_e}\left(p_{01}p_{10} + p_{02}p_{20} + p_{03}p_{30}\right) = \frac{1+3+5}{8m_eL^2}h^2 = \frac{3^2h^2}{8m_eL^2} \tag{31}$$

which is the Schrödinger energy of a free electron on the level $n = 3$.

The procedure can be readily extended to an arbitrary quantum level *n*.

6. Some Special Statistical Behaviour of the Electron Energy Quanta Present in a One-Dimensional Potential Box

Till the present point of our considerations we neglected the properties of the electron statistics applied to the electron gas. In fact the problem of the electron spin, and the Pauli exclusion principle connected with it, were not developed enough at the time of an early competition of the Heisenberg and Schrödinger theories. In principle both the boson and fermion statistics seem to be here applicable, first because of the electron oscillation waves considered in the Heisenberg picture, next because of the double spin-dependent occupation of the energy levels connected with the Schrödinger electron gas state. The fermion-like behaviour of electrons seems however to predominate and our task is to make

only a supplementary insight into the Fermi statistical distribution considered before.

Our point concerns the question whether the highest occupied Fermi level $n_{\max}$ in the one-dimensional gas should be considered as identical with the Fermi energy, or it does represent a distinct energy value. By assuming the second point of view, the Fermi energy—in accordance with former investigations [11] [12]—should be considered as an inflexion point on the Fermi distribution function $F(E)$ plotted as a function of the electron energy *E*.

The function $F(E)$—as it is well known—depends also on the absolute temperature parameter *T*:

$$F(E)=\frac{1}{\mathrm{e}^{-E/kT}+1}. \tag{32}$$

Let us assume *E* to be an abbreviation of a small value of the difference between the Fermi energy E_F and the electron energy on the highest occupied level $n_{\max}$ in the gas which is

$$E(n_{\max})=\frac{h^2 n_{\max}^2}{8m_e L^2} \tag{33}$$

so

$$E=\Delta E^{(1)}=E_F^{(1)}-E^{(1)}(n_{\max})>0. \tag{34}$$

The superscript (1) indicates that—for simplicity—only the gas having a single electron on each of its energy levels is considered. In principle we assume that *E* in (34) is a small number.

The first derivative of $F(E)$ in (32) calculated with respect to the energy *E* gives

$$\frac{\mathrm{d}F}{\mathrm{d}E}=\frac{-1}{\left(\mathrm{e}^{-E/kT}+1\right)^2}\left(-\frac{1}{kT}\right)\mathrm{e}^{-E/kT}=\frac{\mathrm{e}^{-E/kT}}{\left(e^{-E/kT}+1\right)^2}\frac{1}{kT}, \tag{35}$$

whereas the second derivative of $F(E)$ is represented by the derivative of the result obtained in (35):

$$\begin{aligned}\frac{\mathrm{d}^2F}{\mathrm{d}E^2}&=\frac{1}{kT}\frac{\mathrm{d}}{\mathrm{d}E}\left[\frac{\mathrm{e}^{-E/kT}}{\left(\mathrm{e}^{-E/kT}+1\right)^2}\right]=\frac{1}{(kT)^2}\left[\frac{2\mathrm{e}^{-2E/kT}}{\left(\mathrm{e}^{-E/kT}+1\right)^3}-\frac{\mathrm{e}^{-E/kT}}{\left(\mathrm{e}^{-E/kT}+1\right)^2}\right]\\&=\frac{1}{(kT)^2}\frac{1}{\left(\mathrm{e}^{-E/kT}+1\right)^3}\left[2\mathrm{e}^{-E/kT}-\mathrm{e}^{-E/kT}\left(\mathrm{e}^{-E/kT}+1\right)\right]\\&=\frac{1}{(kT)^2}\frac{\mathrm{e}^{-E/kT}}{\left(\mathrm{e}^{-E/kT}+1\right)^3}\left[2-\mathrm{e}^{-E/kT}-1\right].\end{aligned} \tag{36}$$

The *E* in the inflexion point should make (36) equal to zero. To attain that it is enough to require the square-bracket to be vanishing

$$\mathrm{e}^{-E/kT}-1=0 \tag{37}$$

which for small E/kT gives the equation

$$1-1+\frac{E}{kT}-\frac{1}{2!}\left(\frac{E}{kT}\right)^2 \cong \left(1-\frac{1}{2!}\frac{E}{kT}\right)\frac{E}{kT}=0. \tag{38}$$

The last equation is satisfied when

$$E=\Delta E^{(1)}=2!kT, \tag{39}$$

so in this case [see (32)]

$$F\left(E=\Delta E^{(1)}=2kT\right)\cong\frac{1}{\frac{1}{e^2}+1} \tag{40}$$

is obtained at the inflexion point.

In the next step we consider a double occupation of the energy levels in the gas by the electrons having an opposite spin. By assuming that

$$E_F^{(2)}=2E_F^{(1)} \tag{41}$$

and putting for E in (32) the expression

$$E=\Delta E^{(2)}=2\Delta E^{(1)}, \tag{42}$$

we obtain the energy E twice as large as $E=\Delta E^{(1)}$.

A substitution of $E=2\Delta E^{(1)}$ instead of $E=\Delta E^{(1)}$ into Equation (38) gives:

$$\left(1-\frac{1}{2!}\frac{2\Delta E^{(1)}}{kT}\right)\frac{2\Delta E^{(1)}}{kT}=0 \tag{43}$$

so

$$\Delta E^{(2)}=2\Delta E^{(1)}=2kT. \tag{44}$$

Therefore the result (44) obtained for a double occupation of the quantum states implies a reduction of $\Delta E^{(1)}$ defined in (39) to a single kT.

Both results for $\Delta E^{(1)}$ and $\Delta E^{(2)}$ vanish at $T=0$ giving respectively

$$E_F^{(1)}=E\left(n_{\max}\right) \tag{45}$$

and

$$E_F^{(2)}=2E\left(n_{\max}\right). \tag{45a}$$

7. Possible Duality of Statistics Applied to the Electron Quantum Levels

A duality of the boson and fermion statistics which can be applied to the electron levels can be detected by examining the energy of the level ensembles obtained in two different ways. For the electron-gas case a better insight seems to be provided by the Schrödinger's method because of its simplicity.

The eigenvalues of the free-electron Hamiltonian considered by Schrödinger (see Sec. 2) are:

$$E_n=\frac{n^2h^2}{8m_eL^2} \tag{46}$$

where

$$n = 1, 2, 3, \cdots \tag{46a}$$

giving the results identical to those obtained in (5). These results can be successfully considered with aid of the Fermi-Dirac statistics.

But another approach than that in (46) and (46a) can be obtained for the electrons in a one-dimensional box when instead of the stationary states of the energy Hamiltonian in (1) a spectrum of energies due to the electron oscillations within the potential box is considered.

Because of a free-electron character of the particles we obtain from (11) and (12):

$$E_n^{\text{osc}} = \frac{nh^2}{4m_e L^2}. \tag{47}$$

Clasically the electron having the velocity v_n undergoes the way $2L$ along the box in course of the time period T_n; see (9) and (10).

Evidently the energy in (47) is by a factor of

$$n/2 \tag{48}$$

smaller than the energy obtained in (5) and (46). This means that $n/2$ oscillators are required to provide the energy equal to a single eigenenergy state labeled by a given n indicating a large degeneracy of the oscillator energies necessary for any large n.

In effect the energy of a one-dimensional electron gas can be considered also as a superposition of a large number of the boson energy quanta due solely to the electron oscillations.

8. Summary

The paper compares two approaches to the energy levels of a free-electron one-dimensional gas done respectively from the point of view of the Heisenberg and Schrödinger quantum theory. This comparison seems to be absent in the literature.

As a starting point we take into account the Schrödinger wave-mechanical calculation which is very simple. In the next step the electrons are considered as oscillators and the Heisenberg matrices are applied. In fact two different kinds of the Heisenberg's Hamiltonian can be examined for free electrons on condition the constant parameters entering the matrices are suitably modified.

A short calculation concerning the position of the Fermi level in the gas as a function of the absolute temperature has been added. It should be noted that the statistics of quantum energy levels presented in both Heisenberg and Schrödinger theories can be different from that valid for the fermions alone.

Conflicts of Interest

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

References

[1] Schrödinger, E. (1926) *Annalen der Physik*, **79**, 361. https://doi.org/10.1002/andp.19263840404

[2] Schrödinger, E. (1926) *Annalen der Physik*, **79**, 489. https://doi.org/10.1002/andp.19263840602

[3] Schrödinger, E. (1926) *Annalen der Physik*, **80**, 437. https://doi.org/10.1002/andp.19263851302

[4] Schrödinger, E. (1926) *Annalen der Physik*, **81**, 109. https://doi.org/10.1002/andp.19263861802

[5] Heisenberg, W. (1925) *Zeitschrift für Physik*, **33**, 879. https://doi.org/10.1007/BF01328377

[6] Van der Waerden, B.L. (1968) Sources of Quantum Mechanics. Dover Publications, New York. https://doi.org/10.1063/1.3035086

[7] Rubinowicz, W. (1957) Quantum Theory of the Atom. 2nd Edition, Polskie Wydawnictwo Naukowe, Warszawa. (In Polish)

[8] Landau, L.D. and Lifshits, E.M. (1972) Quantum Mechanics. Izdatelstvo Nauka, Moscow. (In Russian)

[9] Schrödinger, E. (1926) *Annalen der Physik*, **79**, 734. https://doi.org/10.1002/andp.19263840804

[10] Eyring, H., Walter, J. and Kimball, G.E. (1957) Quantum Chemistry. 2nd Edition, Wiley, New York.

[11] Slater, J.C. (1967) Quantum Theory of Molecules and Solids. Vol. 3, McGraw-Hill, New York.

[12] Ziman, J.M. (1972) Principles of the Theory of Solids. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139644075

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