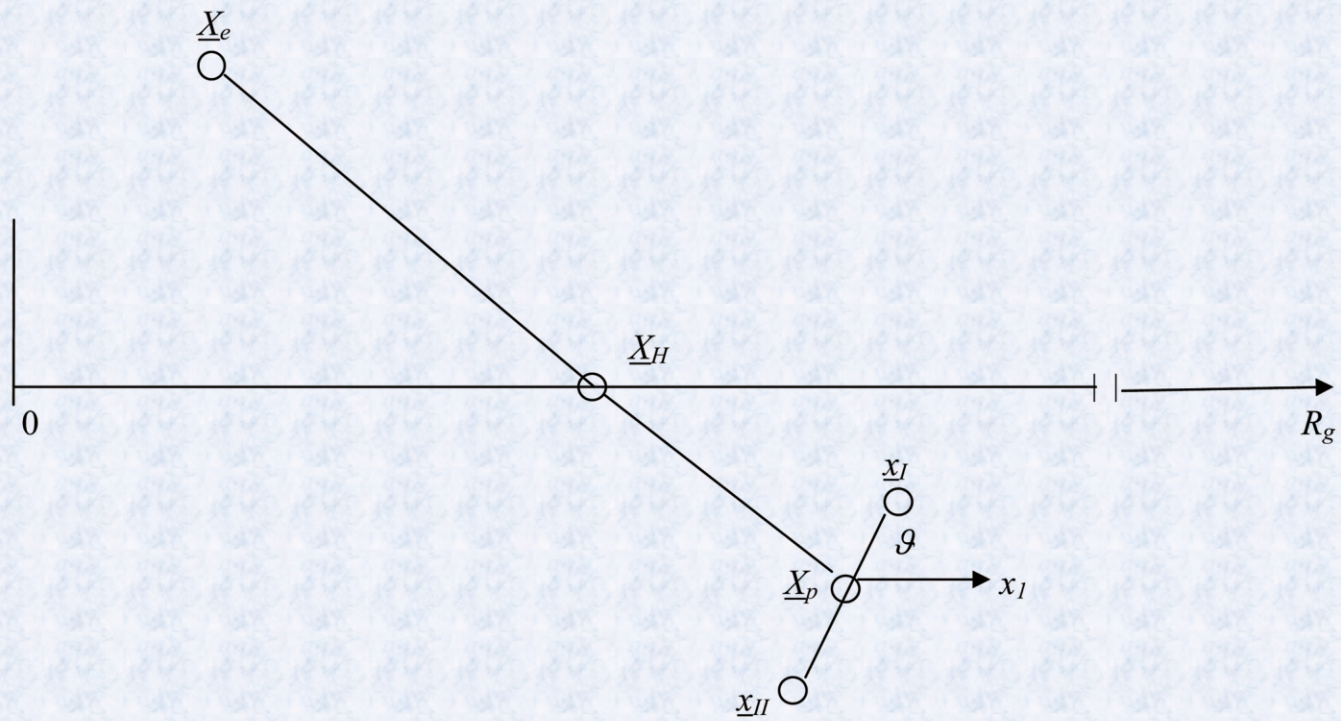
ISSN: 2153-1196 Volume 11, Number 7, July 2020

Scientific Research Publishing

Journal of Modern Physics



ISSN: 2153-1196
9 772153 119007 07

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

Journal Editorial Board

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

Editor-in-Chief

Editorial Board

Scientific Research Publishing

Journal of Modern Physics, 2020, 11, 967-1122
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Table of Contents

Volume 11 Number 7 July 2020

The figure on the front cover is from the article published in Journal of Modern Physics, 2020, Vol. 11, No. 7, pp. 967-975 by F. C. Hoh.

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, 967-975
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

On the Ratios Dark Matter (Energy)/Ordinary Matter ≈ 5.4(13.6) in the Universe

F. C. Hoh

Retired, Dragarbrunnsg. 55C, Uppsala, Sweden
Email: hoh@telia.com

How to cite this paper: Hoh, F.C. (2020) On the Ratios Dark Matter (Energy)/Ordinary Matter ≈ 5.4(13.6) in the Universe. *Journal of Modern Physics*, **11**, 967-975.
https://doi.org/10.4236/jmp.2020.117060

Received: June 5, 2020
Accepted: June 27, 2020
Published: June 30, 2020

Abstract

An upper limit of the average ratio dark matter/ordinary matter in galaxies is estimated to be 8.4, in agreement with the observed ratio 5.4. Upper limit of the average ratio dark energy/ordinary matter for slowly moving protons in the outer parts of the universe is estimated to be 8.4, much less than the observed ratio 13.6. The discrepancy is tentatively attributed to that the bulk of the protons in these outer parts of the universe moves fastly and their contribution to dark energy has not been estimated. The positive and negative relative energies between the diquark and quark in the proton play the roles of dark energy and dark matter, respectively.

Keywords

Relative Energy Between Quarks, Scalar Strong Interaction Hadron Theory, Ratio of Dark Matter to Ordinary Matter, Ratio of Dark Energy to Ordinary Matter, Gravitational Pull of Diquark and Quark, Proton Orbit, Diquark Equations

1. Introduction

The mass-energy density of the dominant cosmic constituents averaged over the entire universe is [1]

$$\text{Dark energy: Dark matter: Ordinary matter} \approx 68\%:27\%:5\% \tag{1.1}$$

Dark energy and dark matter has been interpreted to be the positive and negative, respectively, relative energy between the diquark *uu* and the quark *d* in the protons in the hydrogen gas permeating the universe [2]. Subsequently [3], the negative relative energy has been used to account for the galaxy rotation curve instead of the unobserved dark matter. The positive relative energy has been used to account for the acceleratingly expanding universe instead of the unobserved dark energy. Further, by including the negative relative energy, a neutron

star with a mass near the Tolman-Oppenheimer-Volkoff limit M_{TOV} ~3 solar masses is prevented from collapsing into a gravitational singularity but becomes a "black neutron star".

These qualitative results however do not shed any light on the observed ratios in (1.1). The purpose of this paper is to extend [3] to provide some quantitative estimates on the limits on these ratios (1.1). The proton size is calculated in Section 2. In Section 3, the hydrogen atom structure is illustrated. Section 4 shows the proton structure and estimated upper limits of the ratio "dark matter"/ordinary matter in galaxies. Analogous estimates of "dark energy"/nonrelativistic ordinary matter in a small part of the outer part of the universe are given. In the **Appendix**, gravitational effects are included. Equations of motion for a diquark are constructed. When combined with the equations of motion for a quark, equations of motion for a diquark-quark baryon obtained earlier are recovered.

2. Proton Wave Functions and Its Size

As in [3] [4], the scalar strong interaction hadron theory (SSI) [4] provides the framework. Parts of Sections 2 - 5 in [3] are reproduced for use here together with an extension. Let a *uu* diquark A at x_I interact with a d quark B at x_{II} via a scalar interaction $V_{AB}(x_I, x_{II})$. Since the quarks are not observable, introduce the proton coordinate X_p and the relative coordinate x,

$$x^\mu = x_{II}^\mu - x_I^\mu, \quad X_p^\mu = (1-a_m)x_I^\mu + a_m x_{II}^\mu \tag{2.1}$$

where a_m is an arbitrary real constant. The laboratory coordinate X_p is an observable but x is not but "hidden". The total proton wave function is of the form [3 (3.2)],

$$\begin{aligned} \chi_{0\dot{b}}(x_I, x_{II}) &= \chi_{0\dot{b}}(\underline{x})\exp\left(-iK_\mu X_p^\mu + i\omega_K x^0\right) \\ \psi_0^a(x_I, x_{II}) &= \psi_0^a(\underline{x})\exp\left(-iK_\mu X_p^\mu + i\omega_K x^0\right) \end{aligned} \tag{2.2}$$

$$K_\mu = (E_K, -\underline{K}) \tag{2.3}$$

Here, a, $\dot{b}$ run from 1 to 2. In the rest frame considered in [4], $\underline{K} = 0$, E_0 is the proton mass and $-\omega_0$ is the relative energy between *uu* and *d*.

Equation (1.3) is governed by the baryon wave equations [3 (2.3)] which contains the derivatives $\partial_I^{a\dot{b}}$ and $\partial_{II\dot{e}f}$ with respect to x_I and x_{II}, respectively, appear. For the present case, these derivatives leads to the same result [3 (5.1)] for

$$a_m = \frac{1}{2} + \frac{\omega_0}{E_0} \tag{2.4}$$

[3 (5.2)]. Since a_m can range from $-\infty$ to $+\infty$, the relative energy $-\omega_0$ can likewise do so. This provides the basis of the results in [3].

$\chi_0(\underline{x})$ and $\psi_0(\underline{x})$ in (1.3) have been decomposed into an angular part and a radial part $g_0(r)$ and $f_0(r)$ in [3 (3.4)]. Here, the diquark-quark distance $r = |\underline{x}|$. $g_0(r)$ and $f_0(r)$ have been solved and plotted for neutron in **Figure 1** in [3]. Since the masses of proton and neutron are almost the same, as do those of the u and d quarks, this plot can be used here. Inspection of this figure led to the estimate that the average $r = r_a \approx 4$ fm which has been employed in [3 Sections 6 and 8].

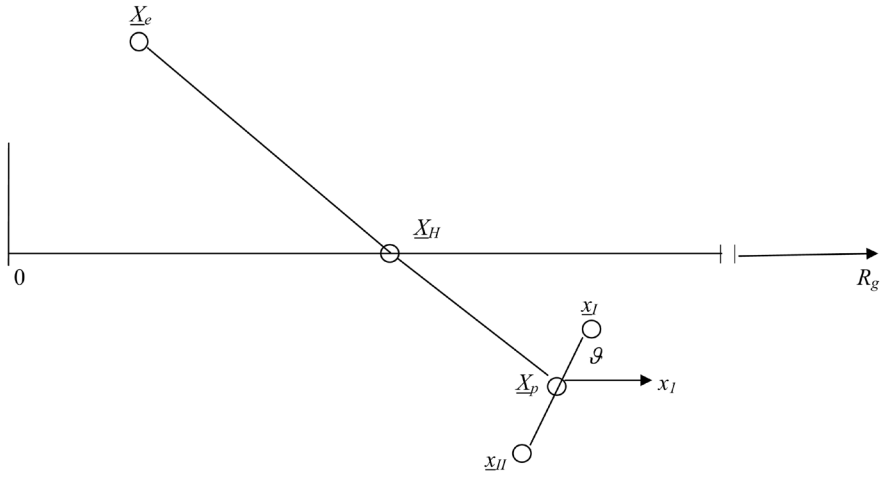


Figure 1. Structure of a rest frame hydrogen atom near the center of a galaxy from which the radius R_g departs. $\underline{X}_e$ denotes the electron coordinate in this laboratory frame, $\underline{X}_p$ that of the proton and $\underline{X}_H$ their center of mass or hydrogen atom. For illustration, the proton-electron mass ratio m_p/m_e has been deliberately lowered to about 2 so that $\underline{X}_e$ falls within the figure. The diquark *uu* is at $\underline{x}_I$ and quark *d* at $\underline{x}_{II}$. A "hidden", relative coordinate system $\underline{x} = \underline{x}_{II} - \underline{x}_I = (x_1, x_2, x_3)$ centered at $\underline{X}_p$ is introduced. ϑ denotes the angles in this "hidden" $\underline{x}$ space and is randomly distributed here.

This r_a value may be improved by weighing r with the squares of $\chi_{0\dot{b}}(\underline{x})$ and $\psi_0^a(\underline{x})$ and averaging over the relative space. With the plots in [3], computation yields for the average $r = |\underline{x}|$

$$\begin{aligned} r_a &= \frac{\int \mathrm{d}\underline{x}^3 r\left(\left(\chi_{0\dot{b}}(\underline{x})\right)^* \chi_{0\dot{b}}(\underline{x}) + \left(\psi_0^a(\underline{x})\right)^* \psi_0^a(\underline{x})\right)}{\int \mathrm{d}\underline{x}^3 \left(\left(\chi_{0\dot{b}}(\underline{x})\right)^* \chi_{0\dot{b}}(\underline{x}) + \left(\psi_0^a(\underline{x})\right)^* \psi_0^a(\underline{x})\right)} \\ &= \frac{\int \mathrm{d}r r^3 \left(g_0^2(r) + f_0^2(r)\right)}{\int \mathrm{d}r r^2 \left(g_0^2(r) + f_0^2(r)\right)} = 3.23 \text{ fm} \end{aligned} \tag{2.5}$$

3. Hydrogen Atom Structure

The structure of the hydrogen atom in the rest frame in SSI is illustrated in **Figure 1**.

The electron cloud is symmetrically concentrated within a sphere with Bohr radius $a_0 \approx 0.53$ Å centered at the center of mass coordinate $\underline{X}_H$. The associated radial wave functions are

$$\begin{aligned} &R_{00}(r_e) \propto \exp(-r_e/a_o), \quad R_{10}(r_e) \propto (2 - r_e/a_o)\exp(-r_e/a_o), \\ &r_e = |\underline{X}_e - \underline{X}_H| \end{aligned} \tag{3.1}$$

Here, $\underline{X}_e$ is the actual (not scaled down as is illustrated in **Figure 1**) electron coordinate in the observable laboratory frame. R_{00} refers to the ground state and R_{10} to the first radically excited state.

Accompanying this electron cloud is an associated proton cloud likewise symmetrically distributed around $\underline{X}_H$. This cloud is analogously concentrated inside a sphere with the proton Bohr radius

$$a_{0p} = a_0 \left(m_p / m_e \right) = 28.8 \text{ fm} \tag{3.2}$$

The observable mass m_p and charge $+e$ of the proton reside in the laboratory frame proton coordinate $\underline{X}_p$ in Figure 1. The diquark uu and quark d are located at the unobservable coordinates $\underline{x}_I$ and $\underline{x}_{II}$, respectively. $\underline{x}_{II} - \underline{x}_I = \underline{x}$ is the "hidden" unobservable relative space "isolated" from the observable laboratory space $\underline{X}$.

4. Upper Limit of Ratio of Dark Matter to Ordinary Matter in Galaxies

In [3 Section 8], dark matter has been used to account for the galaxy rotation curve qualitatively. Defining the ratio of dark energy/ordinary matter to be R_{DE} and dark matter/ordinary matter to be $-R_{DM}$, (1.1) yields

$$\begin{bmatrix} R_{DE} \\ R_{DM} \end{bmatrix} = -\frac{\omega_0}{E_0} = \begin{bmatrix} 68\%/5\% = 13.6 & \text{for dark energy} \\ -27\%/5\% = -5.4 & \text{for dark matter} \end{bmatrix} \tag{4.1}$$

As the galaxy in Section 3 expands, the hydrogen atom in Figure 1 will move outwards from the galaxy center and becomes far away from it. This has been considered in [3 Section 8]. The constituents of this atom are illustrated in Figure 2.

The outward motion of the atom is due to thermal expansion of the young galaxy containing it and proceeds via Coulomb force between the electrons in the hydrogen gas; the proton gets dragged along. The gravitational pull from matter in the central part of the galaxy on the proton produces a negligible shift of X_p and asymmetry in the proton cloud due to the strong restraining electron-proton Coulomb force. Since the quarks are in the "hidden" relative space and do not interact with the electron in the laboratory space but interact with the ambient gravitational field [3 V_{BG} in (2.1)], [2 end of Section 5], they will be pulled back from X_p by these gravitational forces and lag behind to the left of X_p. Further, the diquark uu at x_I is about twice as heavy as the quark d at x_{II} [4 Table 5.2] and will feel twice the gravitational pull and hence lies closer to the galactic center and to the left of x_{II}, also mentioned below (A6).

Since the proton is largely confined inside the proton Bohr sphere, its constituents, the diquark uu at $\underline{x}_I$ and quark d at $\underline{x}_{II}$, are likewise confined to the same

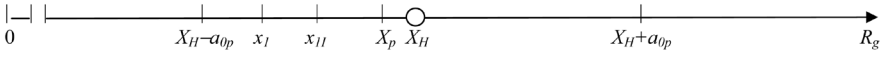


Figure 2. It develops to the right part of Figure 3 in [3]. It illustrates the constituents of the hydrogen atom in Figure 1 moved outwards from the galaxy center to a region far from it. The proton Bohr orbit with radius $a_{0p} \approx 28.8$ fm corresponds to the electron orbit with the Bohr radius $a_0 \approx 0.53$ Å in (3.1). Both are centered at the center of mass X_H which is also the center of the hydrogen atom. The proton cloud is concentrated inside a sphere of radius a_{0p}. This sphere intersects the R_g axis at $X_H \pm a_{0p}$. Here an X_p point to the left of X_H is chosen for consideration; the associated X_e lies far to the right and is not shown. The associated diquark uu is located at x_I and quark d at x_{II} closer to the galactic center $R_g = 0$ due to differential gravitational pull.

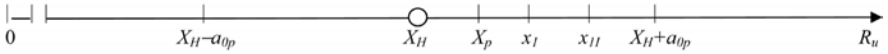


Figure 3. It corresponds to the right part of Figure 4 in [3]. It illustrates the constituents of the hydrogen atom in **Figure 2** moved outwards to an outer region of the universe which is beginning to expand acceleratingly. The horizontal line represents the distance R_u from some unspecified inner region of the universe. The figure has the same form as **Figure 2** but with quarks moved to the right side of the proton coordinate $\underline{X}_p$. The associated $\underline{X}_e$ lies far to the left and is not shown.

proton cloud sphere. This provides a limit $\underline{x}_I \geq X_H - a_{0p}$ in **Figure 2**. The above mentioned symmetry yields an average value $X_p = X_H$. Further, the mean value of $|x_{II} - x_I| = |\underline{x}| = r_a \approx 3.23\,\text{fm}$ in (2.5). Inserting these three relations into (2.1), (2.4) and (4.1) leads to the estimated limit

$$R_{DM} \geq \frac{1}{2} - a_m = \frac{1}{2} - \frac{a_{op}}{r_a} = -8.4 \tag{4.2}$$

consistent with the observed average of −5.4 in (4.1) which lies near the middle of the allowed range 0 to −8.4..

5. Ratios of Dark Energy to Nonrelativistic Ordinary Matter in Outer Parts of Universe

As the hydrogen atom in **Figure 2** continues to move outwards, it enters the outer regions of the universe, as has been considered in [3 Section 9]. In the beginning of this region, the left part of Figure 4 in [3] shows that this hydrogen atom, being too far away from the galaxy's central region, is no longer affected by any gravitational force and reverts from the configuration of **Figure 2** back to that in **Figure 1**. This configuration is unstable with respect to dark energy generated and eventually turns into the configuration on the right part of Figure 4 in [3] where the quarks in the proton are being pushed outwards acceleratingly by the increasing dark energy generated in this region. The situation is illustrated in **Figure 3**.

In this region, the protons are expected to move fastly so that the above treatment, valid for rest frame and approximately also for slowly moving protons, no longer holds. The rest frame treatment in [4 Ch.10-12] no longer holds and no result can be reached. However, just before this region is reached, the norelativistic left part of **Figure 3** in [3] is approximately applicable so that a treatment analogous to that in section 4 can be carried out. Applying **Figure 3** instead of **Figure 2** with the criterion $\underline{x}_I \geq X_H - a_{0p}$ above (4.2) replaced by $\underline{x}_{II} \leq X_H + a_{0p}$ turns (4.1) to

$$R_{DE} \leq \frac{1}{2} - a_m = \frac{1}{2} + \frac{a_{op} - r_a}{r_a} = 8.4 \tag{5.1}$$

This disagrees with the observed $R_{DE} = 13.6$ in (4.1). This is not surprising in view of the consideration above (5.1); (5.1) holds only for slowly moving protons which constitute only a small part of the protons in **Figure 3**. The bulk of the protons there moves fastly and the rest frame treatment here does not hold.

Conflicts of Interest

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

References

[1] Wikipedia (2020) https://en.wikipedia.org/wiki/Main_Page

[2] Hoh, F.C. (2019) *Journal of Modern Physics*, **10**, 635-640. https://doi.org/10.4236/jmp.2019.106045

[3] Hoh, F.C. (2019) *Journal of Modern Physics*, **10**, 1645-1658. https://doi.org/10.4236/jmp.2019.1014108

[4] Hoh, F.C. (2019) Scalar Strong Interaction Hadron Theory. II Nova Science Publishers, New York.

Appendix. Inclusion of Gravitation and Equation of Motion for Diquark

The theory SSI [4] underlying the above results does not include gravitation. These results have been obtained by phenomenologically joining gravitation to the baryon sector of SSI. Here, gravitation will be introduced into the equations of motion for quarks, from which a set of somewhat generalized equations of motion for baryons will be constructed. This construction differs somewhat from the one step construction in [4 Ch 9] in that it is divided into two steps. In the first one, equations of motion for diquark are written down. These are then combined with a set of quark equations to form the baryon equations. This is done because diquark plays a decisive role in the above sections.

For the present applications, it is sufficient to introduce a scalar gravitational potential into the quark equations, such $V_{GB}(x_{II})$ in [3 (2.1)],

$$V_{GB}\left(x_{II}\right)=m_B\varepsilon_{II},\quad \varepsilon_{II}=\varepsilon_{II}\left(\left|\underline{x}_{II}\right|\right)=-\frac{GM_g}{\left|\underline{x}_{II}\right|} \tag{A1}$$

Here, G is the gravitational constant, m_B the mass of the d quark B and M_g the mass of the galaxy containing quark B. $\underline{x}_{II}$ is the position of this quark in a coordinate system with its origin at the center of the galaxy denoted by 0 in **Figure 2** above. ε_{II} is typically a small number of the magnitude of $!0^{-6}$ in the middle of the Milky Way. The gravitational force on this quark is thus negligible but not relative to that on the accompanying diquark in (A6) below.

The potential (A1) containing m_B can be moved to the right side of [3 (2.1)] which now reads

$$\partial_{II}^{d\dot{e}}\chi_{B\dot{e}}\left(x_{II}\right)-i\left(V_{BC}\left(x_{II}\right)+V_{BA}\left(x_{II}\right)\right)\psi_B^d\left(x_{II}\right)=im_B\left(1+\varepsilon_{II}\right)\psi_B^d\left(x_{II}\right) \tag{A2.a}$$

$$\partial_{II\,\dot{e}f}\psi_B^f\left(x_{II}\right)-i\left(V_{BC}\left(x_{II}\right)+V_{BA}\left(x_{II}\right)\right)\chi_{B\dot{e}}\left(x_{II}\right)=im_B\left(1+\varepsilon_{II}\right)\chi_{B\dot{e}}\left(x_{II}\right) \tag{A2.b}$$

which has the same form as the starting [4 (9.1.2a, 2b)] with m_B multiplied by $(1+\varepsilon_{II})$.

Potentials corresponding to (A1) are now introduced into the quark equations [4 (9.1.1, 3)] for quark A at x_I and quark C at x_{III}, respectively. These now read

$$\partial_I^{a\dot{b}}\chi_{A\dot{b}}\left(I\right)-i\left(V_{AB}\left(I\right)+V_{AC}\left(I\right)\right)\psi_A^a\left(I\right)=im_A\left(1+\varepsilon_I\right)\psi_A^a\left(I\right) \tag{A3a}$$

$$\partial_{I\,\dot{b}c}\psi_A^c\left(I\right)-i\left(V_{AB}\left(I\right)+V_{AC}\left(I\right)\right)\chi_{A\dot{b}}\left(I\right)=im_A\left(1+\varepsilon_I\right)\chi_{A\dot{b}}\left(I\right) \tag{A3b}$$

$$\partial_{III}^{g\dot{h}}\chi_{C\dot{h}}\left(III\right)-i\left(V_{CA}\left(III\right)+V_{CB}\left(III\right)\right)\psi_C^g\left(III\right)=im_C\left(1+\varepsilon_{III}\right)\psi_C^g\left(III\right) \tag{A4a}$$

$$\partial_{III\,\dot{h}k}\psi_C^k\left(III\right)-i\left(V_{CA}\left(III\right)+V_{CB}\left(III\right)\right)\chi_{C\dot{h}}\left(III\right)=im_C\left(1+\varepsilon_{III}\right)\chi_{C\dot{h}}\left(III\right) \tag{A4b}$$

Following [4 Section 9.2], the V's on the left sides are moved to the right sides and (A3) and (A4) are multiplied together. While wave functions of three quarks were multiplied together and generalized to baryon wave functions in [4], only two quarks A and C are involved here. Analogous to [4 (2.2.2)], the diquark wave functions are

$$\chi_{A\dot{b}}\left(x_I\right)\chi_{C\dot{h}}\left(x_{III}\right)\to\chi_{\{\dot{b}\dot{h}\}}\left(x_I,x_{III}\right)\to\chi_{\{\dot{b}\dot{h}\}}\left(x_I\right) \tag{A5a}$$

$$\psi_A^c(x_I)\psi_C^k(x_{III}) \to \psi^{\{ck\}}(x_I, x_{III}) \to \psi^{\{ck\}}(x_I) \tag{A5b}$$

Here, quarks A and C belong to the same species and are merged together via $C \to A$ and $x_{III} \to x_I$. This leads to $V_{AC} = V_{CA} = 0$. The multiplication and (A5) yields the diquark wave equations

$$\partial_I^{a\dot{b}} \partial_I^{g\dot{h}} \chi_{A\{\dot{b}\dot{h}\}}(x_I) = -\left(V_{AB}^2(\underline{x}_I) + m_A^2(1+\varepsilon_I)^2\right)\psi_A^{\{ag\}}(x_I) \tag{A6a}$$

$$\partial_{I\dot{b}c}\partial_{I\dot{h}k}\psi_{A\{\dot{b}\dot{h}\}}(x_I) = -\left(V_{AB}^2(\underline{x}_I) + m_A^2(1+\varepsilon_I)^2\right)\chi_{A\{\dot{b}\dot{h}\}} \tag{A6b}$$

in which products not formimg diquark have been dropped. Equations (A5) has the same form as the meson wave equations [4 (2.3.22)] in which the quark masses are m_p+m_r in [4 (2.3.27)]. Analogously, the diquark mass here is $2m_A$. The small gravitational pull on the diquark is thus about twice that on the quark mentioned below (A1). This small difference causes, over long time and large distances, the diquark coordinate x_I to lie closer to the galactic center than does the quark coordinate x_{II} in **Figure 2**, as was mentioned there. How this is accomplished has not been investigated.

The potentials equations [4 (9.1.4a, 4c)] are likewise multiplied together to yield

$$\begin{aligned}\overline{\square}_I \overline{\square}_I V_{AB}^2(\underline{x}_I) = \frac{1}{4} g_s^4 \{&\psi_A^{\{ag\}}(x_I)\chi_{A\{ag\}}(x_I) + \psi_A^{\{\dot{a}\dot{g}\}}(x_I)\chi_{A\{\dot{a}\dot{g}\}}(x_I) \\ &+ \psi_B^{\{eh\}}(x_I)\chi_{B\{eh\}}(x_I) + \psi_B^{\{\dot{e}\dot{h}\}}(x_I)\chi_{B\{\dot{e}\dot{h}\}}(x_I)\}\end{aligned} \tag{A7}$$

analogous to [4 (9.2.2, 2.9, 2.11] involving three quarks. The corresponding expression for quark B [4 (9.4.1b)] reads here

$$\overline{\square}_{II} V_{BA}(\underline{x}_{II}) = \frac{1}{2} g_s^2 \{\psi_A^a(x_{II})\chi_{Aa}(x_{II}) + \psi_A^{\dot{a}}(x_{II})\chi_{A\dot{a}}(x_{II})\} \tag{A8}$$

Multiplying together the diquark A equations (A6, A7) and the quark B equations (A2), with the potentials V moved to the right side, and generalized the product wave functions to nonseparable baryon wave functions similar to [4 (9.2.12)] leads to

$$\begin{aligned}&\partial_I^{a\dot{b}}\partial_I^{g\dot{h}}\partial_{II\dot{e}f}\chi_{\{\dot{b}\dot{h}\}}^{\ f}(x_I, x_{II}) \\ &= -i\left(m_A^2 m_B(1+\varepsilon_I)^2(1+\varepsilon_{II}) + \Phi_b(\underline{x}_I, \underline{x}_{II})\right)\psi_{\dot{e}}^{\{ag\}}(x_I, x_{II}) \\ &\partial_{I\dot{b}c}\partial_{I\dot{h}k}\partial_{II}^{d\dot{e}}\psi_{\ \dot{e}}^{ck}(x_I, x_{II}) \\ &= -i\left(m_A^2 m_B(1+\varepsilon_I)^2(1+\varepsilon_{II}) + \Phi_b(\underline{x}_I, \underline{x}_{II})\right)\chi_{\dot{b}\dot{h}}^{\ d}(x_I, x_{II})\end{aligned} \tag{A9}$$

which reduces to the diquark-quark baryon wave equations [4 (9.2.14)] for $\varepsilon \to 0$. The generalization of quark masses in (A9) to internal mass operators [4 (9.3.8, 13, 14)] operating on internal functions [4 (9.3.1)] is unaffected by inclusion of gravitation.

Multiplying together (A7) and (A8) and using the generalization [4 (9.2.2)] reproduces the baryon potential equation [4 (9.2.11)],

$$\overline{\underline{\quad}}_I \overline{\underline{\quad}}_I \overline{\underline{\quad}}_{II} \Phi_b\left(\underline{x}_I, \underline{x}_{II}\right) = \frac{1}{4} g_s^6 \left\{ \chi_{\{\dot{b}\dot{h}\}}^{f}\left(x_I, x_{II}\right) \psi^{\{\dot{b}\dot{h}\}}_{f}\left(x_I, x_{II}\right) + c.c. \right\} \quad \text{(A10)}$$

Journal of Modern Physics, 2020, 11, 976-995
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Representable Model of a Finished and Unlimited 3D Cosmos

Marc Mignonat

Société d'Astronomie des Pyrénées Occidentales, Pau, France
Email: mmignonat@libertysurf.fr

How to cite this paper: Mignonat, M. (2020) Representable Model of a Finished and Unlimited 3D Cosmos. *Journal of Modern Physics*, **11**, 976-995.
https://doi.org/10.4236/jmp.2020.117061

Received: April 28, 2020
Accepted: July 5, 2020
Published: July 8, 2020

Abstract

A model of a 3D unlimited and finished space is presented with the philosophic prejudgement that a physical space must be representable and cannot be a virtual mathematical abstraction. This representable 3D space has 3 radii of curvature and so, is multiconnected as predicted by the theorem of perelman-Poincaré. This model respects the basic principles of the physics (Occam, Maupertuis, Mach, ...) and does not question much of the content of the other models. Another way of seeing is given because: 1) this model avoids the problem of the 2 infinities; 2) it gives an additional explanation to the expansion and to the value of the density always near the critical density; the homogeneity of the cosmos is easier to explain; 3) the attraction is always attractive and it gives an explanation to the measure of the acceleration of the expansion estimated about 6 - 7 billion years and to the great attractor; 4) it predicts the existence of many "ghosts" images, as the illusion of galaxies is older than Big Bang and these galaxies are more evolved when they are older, or for example a greater number of galaxies at a distance of about 2100 Mpc, what can be verified from the NASA/IPAC measurements of 348 galaxies of redshift $v > 1/8c$. Other deductions are verifiable, which should invalidate or confirm this model. In **Appendix**, a mathematical development deliberately simple is made to remain in a representable reality and locate any point in this space. This development can allow to make links with the spaces of Minkowski.

Keywords

Cosmology: Theory

1. Introduction

The purpose is to describe with curvilinear coordinates a finished and thus representable 3D space. In essence, an infinite space is not representable and from

a philosophic point of view, a not representable entity would make us go out of the domain of the physics. Penrose [1] gives as conditions that a theory is valid if the two infinities are eliminated: the infinity of the point of origin and the outside infinity.

The opposite philosophic idea that a physical reality is not representable was remarkably developed at the atomic level by Bohr [2]. In opposition with Kant who separates the subject objector and the observed object, he considers that the object is connected to the subject. The wave or corpuscular aspects are a function of the way of observing of the subject. He is in the continuity of Lamé and Maxwell. Lamé invented curvilinear coordinates in a three-dimensional Euclidean space [3] (he assumed that the equation of Laplace is equal to zero $\Delta V = 0$, and so, he described families of curves and surfaces). But we are always inside an infinite Cartesian space.

For Maxwell and his successors, the photon, described by 2div and 2 rot with infinities, becomes an abstraction and so, is not representable.

This idea that the physics is mathematical is a principle given, among others, by Lamé, Maxwell, Bohr and recently recalled by Tegmark [4] (for him an equation creates reality). This postulate goes against the scientific method of Descartes and against the English empiricists where all start from the observation.

Before, Gauss [5] began to define coordinates on the surface of a sphere. The generalized coordinates of Gauss describe an infinite space; they allow to pass in a space with a dimension more than 2, but, we go then into a space which becomes abstract, purely mathematical, infinite and thus non-representable. So the generalization to 3 dimensions of the formula $\cos s = \cos x \cdot \cos y$ applicable to the surface of a sphere, by considering that s and z are on the surface of another virtual sphere, would give a formula of type $\cos s' = \cos x \cdot \cos y \cdot \cos z$. This formula would allow mathematical developments but does not suit us for the philosophic reason that we do not want to create a non-representable mathematical virtual world, but to stay in the physics, *i.e.* in a representable concrete world.

It is the philosophic position where the physics represents a reality and where the mathematics is at the service of the physics and not the opposite. Einstein and Broglie are from this current. Sommerfeld in 1923 [6], considers that "the classic undulatory theory was not replaced by a better theory and especially for the spectroscopy". For Poincaré, the mathematics is inductive and not deductive (they allow a generalization but remain not predictive) [7].

This superiority of the mathematics permeates the thought of a large number of current physicists and sometimes, unconsciously. Many cosmologists apply to the study of the cosmos, a philosophy which can be relevant but which is debatable as any philosophy. This one was established by Bohr, at the particulate level and not at the macroscopic level (we can note that it is a problem encountered by the mathematics of the general relativity (GR) when for example, we generalize Riemannian manifolds or the coefficients of Christoffel from the dimension 2 to a superior dimension). Poincaré said: "any generalization is a leap into the

unknown." [7]

2. Mathematical, Astronomical and Physical Aspects

2.1. Mathematical Aspect

2.1.1. General Mathematical Aspect of This Type of Space

We take an orthonormalized space in R3 defined by 3 basic curvilinear coordinates ***i, j, k***. Every curvilinear coordinate ***i, j, k***, has a bend radius. Our goal being to present a simple astronomical model and not to make a long mathematical development, we set the same radius of curvature at the three curvilinear coordinates ***i, j, k***. (This condition allows to draw spherical type spaces and simply avoids spheroidal spaces in "rugby ball" more or less crushed). So, we are in a non-Euclidean space of Riemannian type (and therefore not in a space of Lobatchevsky where the radius of curvature is variable.), which allows us to respect the theorem of Lie Poincaré [7]. (This theorem imposes that in a physical space and not mathematical, it is necessary that the space be at "constant curvature" because a figure "flexible and inextensible in motion would then be deformed and therefore unable to move". "only the movement of an invariable figure is possible". This argument by Lie Poincaré eliminates the hyperbolic spaces with negative curvature since in this case. The figure is distorted when it moves...).

The attachment of this radius leads to define a fourth dimension of length. (with 3 different curvature radii, a 5th and 6th dimension of length would be added).

As the curvilinear coordinates ***i, j, k*** are not taken as rectilinear vectors but as coordinates having a radius of curvature *r*, a point following a coordinate will describe a circle and comes back to its starting point, which has the consequence that this space defined on R3 becomes finite (**Figure 1**) (3 curvilinear coordinates is a sine qua non condition to have a finite space).

We can note that the same point can be located by several groups of curvilinear coordinates (The mathematical development is in **Appendix**).

The 2 coordinates ***i, j*** will allow to define a 2D surface. The 3rd curve coordinate ***k*** will define the 3D space. With 3 orthonormated coordinates ***i, j, k*** with the same curvature, 2 cases of figure (sphere or torus) will appear depending on whether the radii of curvature ***i*** and ***j*** are on the same sphere or not. (**Figure 2**) But, in all cases, there is a central zone that does not belong to the 3D space. The 3D space thus defined, is always multiconnex according to the theorem of Perelman-Poincaré which imposes that a 3D space is necessarily multiconnex.

2.1.2. Aspect in Sphere

It is a sphere of radius 2.414*r* (**Figure 3**) with an inaccessible central "hole" and therefore not belonging to the space defined by the coordinates. This "hole" has a radius of 0.414*r*.

The curvilinear coordinate ***k*** always goes perpendicular to the surface defined by the curvilinear coordinates ***i*** and ***j***, which leads to a multiconnex "hypersphere"

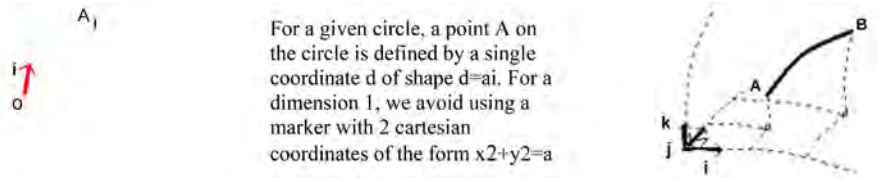


Figure 1. A 3D finite space can be defined by 3 curvilinear coordinates.

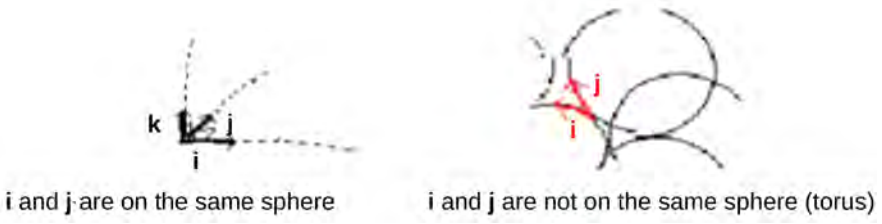


Figure 2. Sphere or torus aspect.

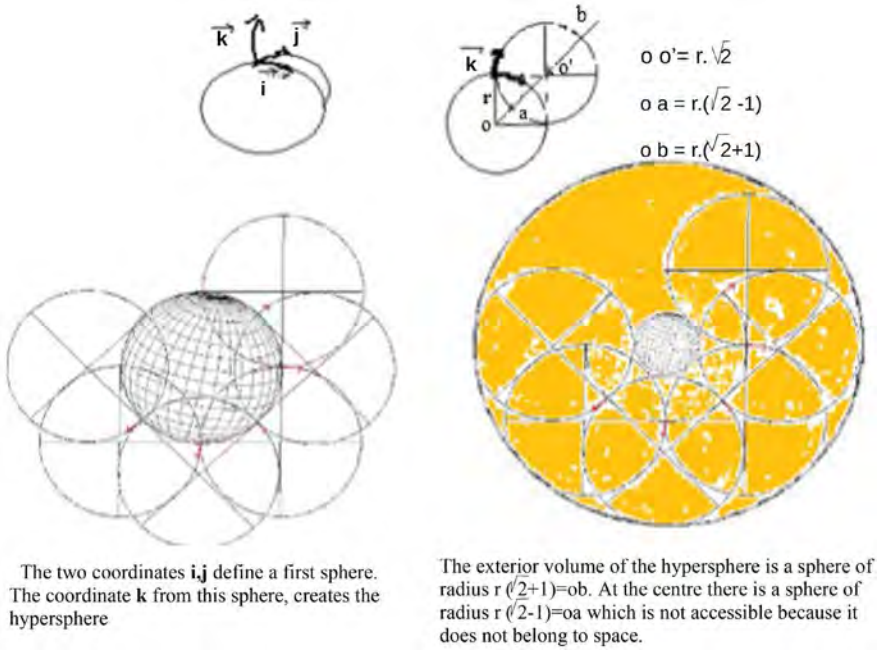


Figure 3. Aspect in sphere.

with for external boundary a sphere of radius $r(2^{1/2} + 1)$ and for internal boundary a sphere of radius $r(2^{1/2} - 1)$. The volume of this space is equal to the volume of the sphere of radius ob decreased by the volume of the sphere of radius oa, or, calculated from the radius r, Vol. = $(56/3)\cdot\pi r^3$

The surface of this space is equal to the surface of the sphere of radius ob increased from the surface of the sphere of radius oa, or calculated from the radius r, Surf. = $24\cdot\pi r^2$.

2.1.3. Aspect in Torus

It is a multiconnex torus of large radius $4.414r$ and small radius $2.414r$ (Figure 4).

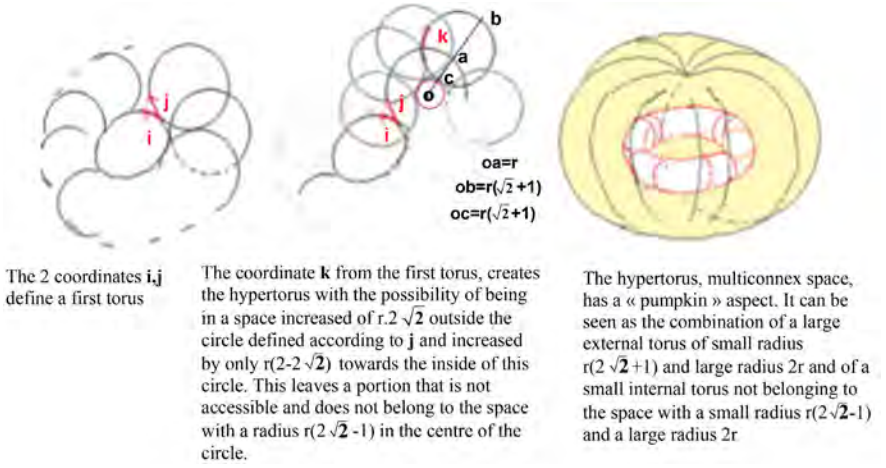


Figure 4. Image of a 3D torus or hypertorus.

The coordinates ***i*** and ***j*** have their curves reversed, which creates the classical 2D torus.

The coordinate ***k*** goes perpendicular to the surface of this torus and thus induces multiconnex 3D hypertorus. The small radius R_p of this hypertore will be $R_p = r(1 + 2^{1/2})$ with r=radius of this space. The large radius R of the 2D torus being equal to $2r$, the torus has the appearance "without hole" or "pumpkin". Inside the hypertorus is a torus not belonging to the space with a small radius $r(2^{1/2} - 1)$ and a large radius $R = 2r$.

The volume V of the 3D torus defined from the radius of curvature r is equal to $V = 16 \cdot 2^{1/2} \cdot \pi r^3$

($V_t = 2\pi r^2 R$; Vol.torus 3D = Vol. torus with small radius $r(1 + 2^{1/2})$ − vol.torus with small radius $r(2^{1/2} - 1)$ with $R = 2r$)

The surface is $S = 16 \cdot 2^{1/2} \cdot \pi r^2$

($S_t = 4\pi rR$; Surf;torus 3D = Surf. Torus small radius $r(1 + 2^{1/2})$ + surf.torus small radius $r(2^{1/2} - 1)$)

2.2. Astronomical Aspect: A Continuous Astronomical Space

As in the 3D sphere (or hypersphere) of Poincaré [8] [9] or the crumpled space of Luminet [10], we find that whatever the direction in which we go, we return to the starting point. The difference is that in the space we describe, the movement can be continuous while there is a discontinuity for the spaces of Poincaré and Luminet. Indeed, Poincaré conceives "modes of conjugation of the faces of polyhedra". He associates by mathematical convention an algebraic surface that forms "cycles of edges" and "cycles of vertices" to another algebraic surface. He notes that these polyhedra are multiconnex. Luminet develops the dodecahedron of Poincaré to make it a model of universe and reminds [11] that "it is infinitely unlikely that the curvature of space is strictly zero..."

On a world map, in the space of Luminet, by imagining the route of a boat, arrived at the edge of the map, we make a discontinuity to extend the journey on another edge of the map. On a globe, as in our space, the motion of the ship, in

imagination, is continuous.

The advantage of being able to include in imagination this four-dimensional space (3-dimensional length and a radius of curvature) in a Cartesian R^3 orthonormated space (Cartesian space without physical reality) is that it can be represented by remaining in a 3-dimensional world.(just as the Riemanian manifolds make it possible to study a space of dimension n while remaining inside this dimension) In the case of spherical aspect, each point can be located by a coordinate of type: $OS^2 = 4R^2\left[\sin^2 w/2\left(1\pm\sin z\right)+\sin^2 z/2\left(1+\sin w\right)\right]$, an event by a formula of type:
$OS^2 = c^2t^2 - 4R_0^2(t)\left[\left(1\pm\sin z\right)\sin^2 w/2+\left(1+\sin w\right)\sin^2 z/2\right]$ with $R = R_0(t)$ (see details of the calculation in **Appendix**).

2.3. Physical Aspect with a Spherical Representation and Curved Coordinates

2.3.1. Spherical Representation

In a physical and not mathematical perspective, the principle of least energy leads to a preference for spherical representation. What we can also translate by the reminder, that isolated, a body seeks to have the maximum volume V for the smallest surface S. In the space, a volume of liquid forms into a ball. Conversely, a red blood cell to have the largest possible gas exchange surface for a given volume will have an toroidal shape.

2.3.2. Consequences to Have Curved Coordinates

We can note that the 4th dimension, "the radius of curvature", if it is associated with a contraction or expansion, so with a change in length over time, could easily be associated with a temporal dimension ($R = R_0(t)$) recalling the Minkowski space. (The above equation defining OS^2 is for a curved space while the Minkowski equation $ds^2 = c^2t^2 - dx^2 - dy^2 - dz^2$ is for a flat space).

In an isolated flat or curved space, If there is an expansion, the orthonormated marker stretches, the distance to be covered ($n \times (\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$) is always the same and the travel time does not change; However, in a curved space, the radius of curvature varies.

We are in a space with curved coordinates R^3 which is not a sub-set of a Cartesian vector space R^3. This set of R^3 encompasses and constitutes the cosmos (there cannot be an orthonormated space R^3 separated from a curved space R^3). In other words, a space cannot be both finite and infinite. If the curved space has a reality, then the Cartesian vector space is purely virtual and has no physical reality of its own since, let us remember, being infinite, by essence, it is unrepresentable; it is only a virtual mathematical tool that allows us to represent real curve space. If one considers that there is a real Cartesian space, then it makes no sense to describe a curved space and the model we present is of no interest. But to assume that there is a real Cartesian space we have to do a philosophical postulate that nothing demonstrates and for which there is no observation.

Thus, without a Cartesian reference outside the system, the travel time to

cross the universe is the same whether the universe is contracted or very dilated, that is to say that the radius of curvature is very small or on the contrary very large. The only variable and measurable element in relation to a measurement standard is the radius of curvature which will have a lower limit (it tends towards zero when expansion grows). The variation of this radius means that there can be no strictly covalent coordinates (the enlarged image of a coordinate is not the same as the radius is different).

If, in imagination, this space with curved coordinates was included in a Cartesian vector space (reference frame outside the system), the travelled distance would not be the same: in a very contracted universe, with an external reference frame, for example, we could see the light making a few cm/s; but in the contracted curved space, these few cm represent the 300,000 km of a more dilated universe; the light still makes, in both cases 300,000 $km \cdot s^{-1}$ (Figure 5).

It can also be said that the ratio between the diameter of an atom and the diameter of a galaxy remains unchanged in both cases. This problem of markers is not or little addressed in the theory of expansion and when we talk about the density of the big bang at its beginning. Without an external reference to the big bang, talking about density makes no sense. To speak about density at the beginning of the big bang presupposes a frame of reference outside the system, therefore presupposes the existence of a previous space. Without an external marker, the Cartesian marker stretches with expansion and the density remains constant (Figure 6).

This problem, that we highlight, does not arise with a system with curved coordinates since the radius of curvature (which will decrease with expansion) constitutes an internal reference to the system allowing to compare volumes.

It should also be noted that there is a central $0.414r$ radius "hole" that does not belong to space. This means that in case of maximum contraction the space cannot be reduced to a mathematical point.

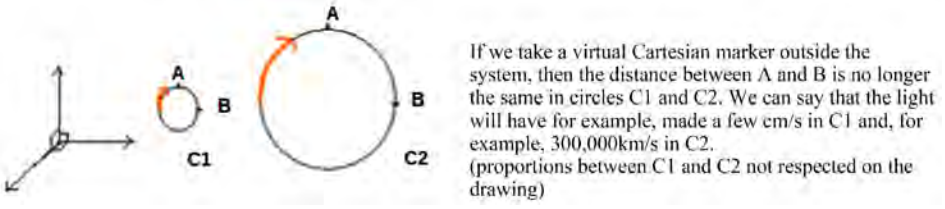


Figure 5. Cartesian marker outside the system.

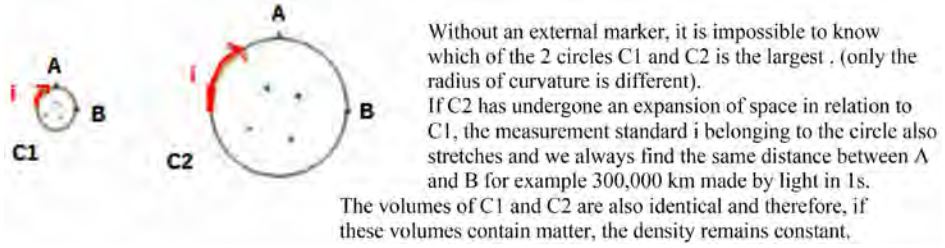


Figure 6. System without an external marker.

3. Cosmological Consequences or Arguments in Favor of This Finite Space

(arguments about representation, multiple images and gravitation are to be seen as theoretical elements since they are verifiable. On GR and black holes, this is only conjectures)

3.1. Problem of Representation

This finite space avoids some of the problems posed by an infinite space, problems which are the same for 1, 2 or 3 dimensions.

To understand the advantage of a finite 3D space, it is necessary to return to the representation of the surface of the Earth which is a 2-dimensional space, finite and unlimited.

Before Anaximandre in antiquity or the arrival of the Jesuits in China, the Earth is a flat disk; a space with 2 dimensions infinite and unlimited since we can always enlarge the plateau or extend a stick beyond the plateau. The understanding of the world is then very complex and calls for insoluble contradictions comparable to our current problem of limits (How do the stars turn and how do they pass under the plateau? What is behind the plateau? Where does the water fall? Does it go into a parallel world? another dimension? a hyperspace or does it pass, as it seems today, by a wormhole to go from west to east?) In ancient Egypt, it was said that the goddess Nout swallowed the sun to the west and gave birth to it in the morning to the east. With our current language, they would have talked about hyperspace and wormhole.

On the surface of the earth, we can find ourselves in 2D; the curvature of the earth (3rd dimension) is only an intellectual construction that we do not use in the system of longitudes and latitudes. Rotondity is a concept that provides a better intellectual understanding of the Earth than a flat Earth.

Similarly, in our curved space where we can walk in 3D; the 4th dimension of curvature is not necessary to locate us, but this 4th dimension is essential to build a finished space.

Of course, it is not because a concept allows for simpler representation that it proves its validity. But reject this concept for this reason is reject the representation allowed by a round Earth and continue to search for how to represent a flat Earth. Now, the simplest representation is the most likely to be true: it is the principle of simplicity (lex parcimoniae) or the principle of the Occam razor.

NB: Note on the traveler of the Middle Ages coming to the end of the world and extending his stick. In our spherical 3D space, it might be pointed out to us that the traveler can always extend his stick, but it is forgetting that here the traveler does not walk on an infinite and unlimited 2D plane but on a 2D sphere; therefore, it will remain in a finite zone and cannot go to infinity; moreover, a stick of astronomical length would bend (**Figure 7**).

3.2. Multiple Images

The 2nd property of a finite space means that a specific event located at a distance

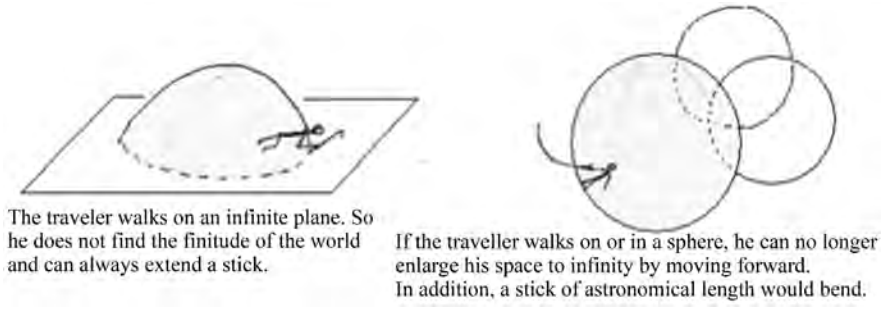


Figure 7. Philosophical problem of the traveller of the middle ages.

x can be perceived at least once, and in an opposite direction at a different distance.

Finite spaces (circle and sphere) have properties that can also be found in 3D spaces such as the Poincaré hypersphere [8], the crumpled spaces of Luminet [10] or the space described above with curved vector markers.

1) Irrespective of the direction we take, we always go back to where we started.

2) The radius of curvature remains the same when we consider only a dimension 2 of the 3D space or a dimension 1 (a circle inscribed in a sphere has the same radius as the sphere).

3.2.1. General Case

In this curved space diagram, whatever the object A, there will always be an image A" in opposition on the circle linking B (observation point) and A.

The shorter the BA' distance, the greater the BA" distance; the total is 14×10^9 light years (assuming a perimeter equal to 14×10^9 LY).

An analogy can be made to sound on Earth. So let us assume that a vector such as sound can be heard throughout the Earth. Let us suppose there was a punctual explosion in London and we were located on the meridian of Greenwich at 1000 km, for example in Tarbes. We will hear the sound in Tarbes 1 hour after the explosion, but also about 39 hours later. (We assume a straight-line path and do not assume that the sound could travel around the Earth several times.) If we do not know that we are on a sphere, but that we believe ourselves on a plane, then we will have recorded 2 explosions located in opposite directions and at distances of 1 hour and 39 hours of journey-sound.

We make here the hypothesis of not deviated light paths, which leads to a spherical geometry with 2 opposite images (**Figure 8**).

A verification would be possible if a galaxy was precisely identified, if we knew how it was going to evolve or if we identified particular characteristics, allowing us to say that this galaxy in another region of the sky is, actually the same galaxy seen at a different age. The idea of 2 images at different ages of the same object can also be found in a particular group or configuration of galaxies that can be more easily predicted mergers, closers, moves away, etc. In the hypothesis of this curved space, it must be considered that all objects located very far away, for example beyond 12 billion LY, are in fact very close to each other (**Figure 9**).

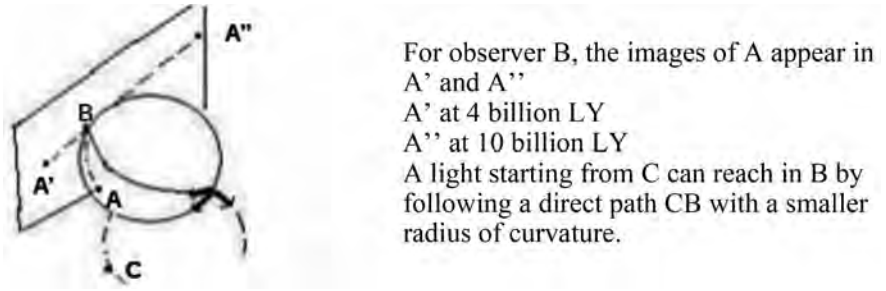


Figure 8. General case with 2 images of the object A.

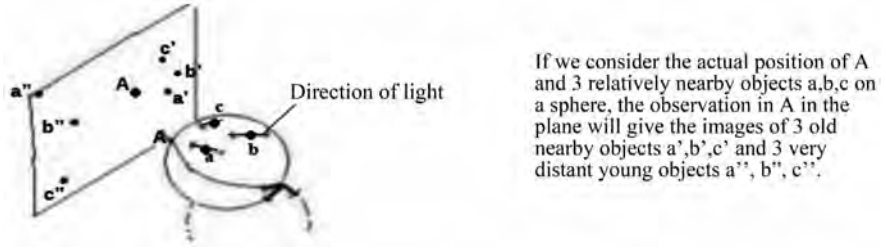


Figure 9. The 3 distant objects a", b", c" are in fact the images of the nearby objects a, b, c.

If we imagine a sphere with a perimeter of 14×10^9 LY, then there is an average deviation of 1˚ every 40×10^6 LY.

NB: If the volume of this space is calculated from a radius $R(1 + 2^{1/2})$, its radius remains equal to R.

3.2.2. Particular Case: When A Is Located at 180˚

(**Figure 10**) in a direction according to the coordinates $\boldsymbol{i}$, $\boldsymbol{j}$ or $\boldsymbol{k}$, in fact the farthest possible point, the images A' of the object A will be located on a celestial circle, All of them are about 7×10^9 years old.

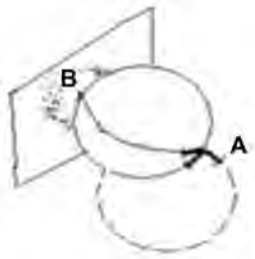
This means that for this distance, the images of a galaxy will be in greater number and a peak of galaxies should be observed at this distance. We took a survey of the 348 galaxies with a $v > 1/8c$, taken from the NASA/IPAC catalogue [12], which seems to confirm this hypothesis (**Figure 11**).

The peak of galaxies is between 2000 and 2500 Mpc and would correspond to a half-turn (median at 2260 Mpc when taking galaxies between 0 and 4500 Mpc). The complete turn would give the time since the big bang either 4500 Mpc +/− 500 Mpc or 14.7 +/− 1.6 billion LY.

This histogram with a second peak may also suggest that the light could have gone more than one turn. But this hypothesis requires a larger statistical series to be verified; in this case, we will have the illusion of perceiving galaxies of age greater than 14 billion years with the apparent paradox that they are all the more evolved as they are old.

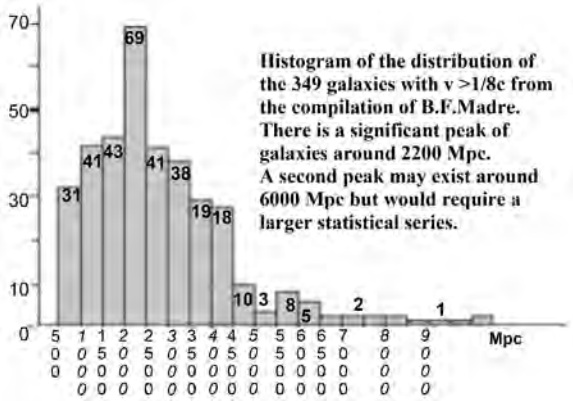
3.2.3. Particular Case at 360˚ (Figure12)

At 4400 Mpc it is an image of ourselves, 14×10^9 years ago. The image can follow all the meridians and make a complete turn to arrive. The image will be found



Light can follow all paths in the sphere of diameter AB. For the B observer, the light from A will form a circle of images in a plane perpendicular to the diameter AB. The light, starting from A on a path which is not on the surface of the sphere of diameter AB, cannot reach point B.

Figure 10. Objects at 180˚.

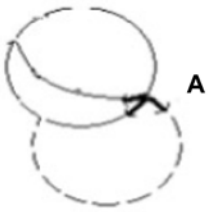


Histogram of the distribution of the 349 galaxies with v >1/8c from the compilation of B.F.Madre. There is a significant peak of galaxies around 2200 Mpc. A second peak may exist around 6000 Mpc but would require a larger statistical series.

3 galaxies around 14,000 Mpc are not reported.

7 galaxies were isolated from the results due to inaccuracy in their distance (2000 Mpc difference between 2 measurements)

Figure 11. Peak of galaxies at 180˚.



The light from A returns to A. All the trajectories being circular, whatever the direction taken, go back to point A. The images from A, for the observer in A, will reach from every point in space and form a celestial sphere.
It is the image of the CMB, from a single point, which comes to us from all directions of space.

Figure 12. Objects at 360˚.

throughout the celestial sphere, so will be identical in all directions (for the cosmic microwave background, the differences of temperature around 3˚K can only be an indication of the hazards encountered by light during its journey and/or images of the different faces of the same object; so, there is no problem in explaining the homogeneity of the universe since every point in the celestial sphere is the image of the same point.). But nothing extraordinary for the theory of the Big bang and it is finally easier to imagine it by recalling that it took place also where we are rather than elsewhere at 4200 Mpc, in an undetermined direction

3.2.4. Beyond 360˚

There is nothing to prevent the light from doing more than one turn and therefore having the illusion of a galaxy older than the big bang with the paradox that the farther away it is, the more evolved it is. (Perhaps this is what the survey from the NED catalogue above suggests with the reservation that the series stu-

died is too small to be significant on the second peak.)

Let us remember that in our curved universe, a large part of the universe is not visible from a given point. Because of the curvature, we cannot see beyond a diameter of 4.5×10^9 LY in a space with a diameter of about 11×10^9 LY. To see beyond 4.5×10^9 LY, you would have to travel. If light can reach us from other points in space (**Figure 7(a)**), there are points from which it cannot reach us (**Figure 9**).

3.3. Gravitation Is Always Attractive, Does Not Need to Be Repulsive

In this curved space, all bodies moving away from each other will, at a given moment get closer without this rapprochement meaning a contraction, the expansion (corresponding to an increase of the radius r) not being prevented from continuing. The analogy with the drift of the continents can be made: (**Figure 13**) we start from the Pangea, the continents drift and then, little by little, get closer. (In the hypothesis of expansion, it is the space that grows between galaxies and not galaxies that move, and we can continue the analogy by saying that it is not the continents that drift but the mantle of Earth that grows from the dorsals).

Because of the attraction that the continents exert on each other, the continents will accelerate their closeness, so as part of the expansion, we will have the impression of an acceleration of the expansion. For example, if we consider the America continent moving away from Europe; America moving closer to Asia, the growing attraction of Asia causes an acceleration of the displacement to the west. Seen from Europe, America seems to be accelerating its distance and one can falsely conclude that the attraction between Europe and America has become repulsive. The speeds of galaxies are increasing! They seem to be moving away, but they are actually getting closer! We have here a hypothesis of explanation of the paradox between expansion on the one hand, acceleration of expansion and galactic moving closer towards a point without mass: "the great attractor" on the other.

Gravitation remains attractive, it's reassuring!

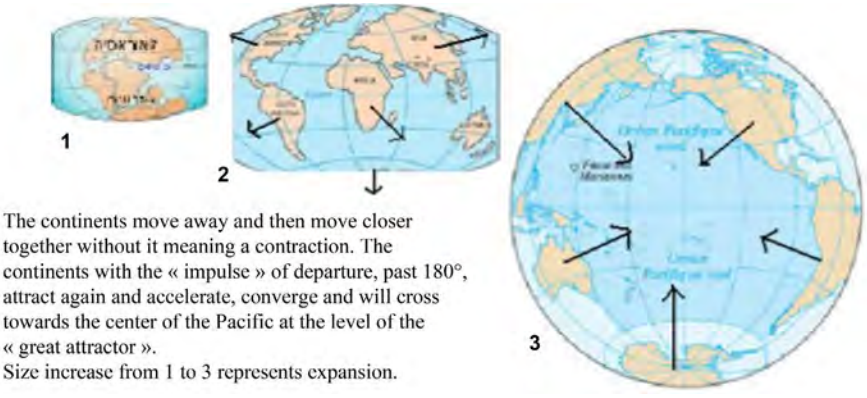


Figure 13. Explanation of acceleration of expansion by gravitation.

Verification: always taking the hypothesis of a 3D sphere of perimeter 14 billion LY, the appearance of the acceleration will be located when 180˚ are traversed, therefore around 7 billion AL. Galaxies will converge to a crossing point at the level of the "great attractor" and it will be necessary to distinguish the acceleration due to the attraction of the masses from the not necessarily increased acceleration of the expansion.

A verification could be seen on the representation of Laniakea [13] where galactic movements are curved and in the direction of an apparently empty point. If the displacements are actually on one or more spheres, then the galaxies face each other and tend to move closer together (see Drawing 3 of **Figure 12**). It is then logical to see them curved trajectories and their accelerations towards the point of convergence can be calculated in a conventional way.

3.4. Conjecture: Does This Space Allows to Give a Representation of the 3D Space Curve of General Relativity? Does a Radius of Curvature Lead to the Existence of a Centrifugal Force?

The surface of the Earth is a curved, finite 2D space. By applying the mathematical formalism found in general relativity (GR), an object moving on its surface cannot leave this surface, follows a geodesic and therefore does not, by definition, undergo centrifugal force. It is the introduction of the 3rd dimension (the terrestrial radius) which, mathematically, allows a displacement in a direction parallel to the radius and thus authorizes the inertial forces (an aircraft follows a geodesic and undergoes centrifugal and Coriolis forces).

The GR mathematically describes a non-representable 3-dimensional curve space. As with the previous 2D space, any object follows a geodesic where by definition there is no centrifugal force. In the curved space we describe, the existence of a radius of curvature (4th dimension) would lead, as in the previous case, that any object moving with a velocity v would undergo a centrifugal force even if it followed a geodesic. The GR is a principle, not in opposition to the Newtonian theory [14] which is neutral on the question of the finite or the infinite of space; the question of a 4th dimension of length, constitutive of space is irrelevant for the GR. It is also necessary to recall that the GR is placed in a 3-dimensional curved space, built by a generalization of a 2-dimensional space. So, it is connex, thus not belonging to the world of physics but only to that of mathematics. Now, the Poincaré conjecture become theorem of Perelman Poincaré previews that "any closed dimension 3 manifold having a finite fundamental group has a spherical geometry", that is to say is covered by the 3-sphere. This space of dimension 3 is therefore necessarily multiconnex.

Logically, the calculation of this centrifugal force, if it existed, would make it possible to know the radius of curvature of the space (Another way to measure the radius would be to follow two different paths between 2 points A and B and measure different gyroscopic orientations).

This hypothetical centrifugal force will be stronger when the radius is smaller and this can then be a hypothesis to the engine of the expansion of the Universe.

So, we arrive at a balance between force of attraction and centrifugal force which would imply that we would tend towards the critical density which would in fact be a density of equilibrium.

If we want to avoid talking about centrifugal force, we can reason from entropy. Entropy S is increased when the matter disperses in a larger volume ($\delta S = -PdV$). But if, at the same time, the volume increases because of the enlargement of the reference mark linked to the expansion, ($R = R_0(t)$ = comobile distances and therefore comobile volume) the measured volume is identical, the density of the matter is identical and the increase of entropy is then zero (Figure 14).

It is not forbidden in the space we describe, that areas are expanding and others locally, such as black holes, are contracting.

In this 3D space we can make the comparison with the surface of the earth which is a 2D surface but embossed between the seabed and the high mountains. Perhaps, the problem of the finiteness of our 3D space is no different from the 2D surface of the Earth. We cannot leave 2D space even if the relief deforms this space or if an expansion or contraction of the Earth takes occur.

3.5. For Black Holes (BH)

The multiconnex space described above is also to be studied within the black hole which is a contraction zone.

When the mass of the BH increases, the radius of Schwarzschild increases, the radius of the BH itself is supposed to decrease and it is accepted that the BH curves more and more the space that is here in contraction. This curvature being made in 3 dimensions, the theorem of Perelman-Poincaré imposes a multiconnex space. If we just assume that the curvature of the space is done according to my orthonormated curved marker in R^3, then it will appear at the center of this hole for sufficient density, an area of radius $R(2^{1/2} - 1)$ not belonging to this space.

This central zone not belonging to space will always exist even if the radius decreases.

This hypothesis of a 3-dimensional curved multiconnex space eliminates the problem of infinite density at one point. However, it does not eliminate the problem of unlimited increasing density but may open the door to simpler conjectures.

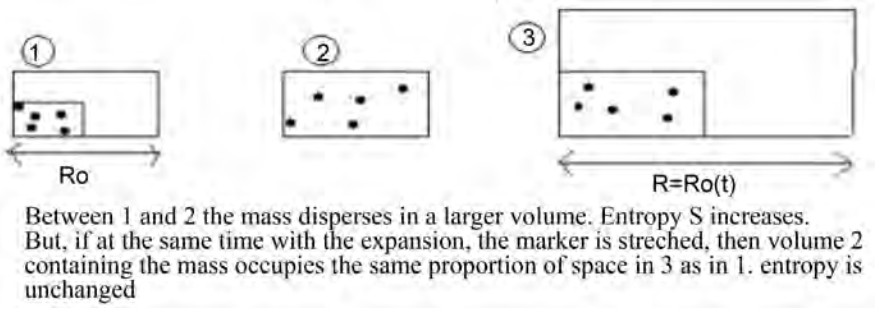


Figure 14. Entropy is unchanged.

4. Hypotheses and Conclusion

What can be said is that, in total, the topological modification would allow the vision of a cosmos that seems to us more coherent with verifiable theoretical consequences:

1) This cosmos respects several principles:

- First, the principle of simplicity (lex parcimoniae or principle of the Occam razor)
- This space, like others, respects the principle of Mach. The observation that there is always an apparent movement implies that there is no centre or particular point (spherical space).

2) This topology meets the theoretical prerequisites of Perelman-Poincaré which requires any space with 3 dimensions of curvature to be multiconnex. And to those of Penrose by removing the problem of the 2 infinite: the infinity of the origin point of the big bang since the central zone does not belong to space; the outer infinity since space is finite. The cosmos is then representable without the mathematical impossibilities or singularities related to the infinite.

The notion of big or small does not make sense if we do not have Cartesian benchmarks outside the system. Only the measurement of the radius of curvature can tell whether this space is with a small or large expansion. The radius of curvature may be the only possible objective measure within this space because it is not dependent on a measurement standard external to the system.

3) The movement in this space is continuous. There is no discontinuity as in the hyperspheres of Poincaré or Luminet where, we come out one side and we come back another side. (Problem of mapping a Round Earth)

4) The homogeneity of the cosmos can be explained by the shape of this space where seemingly distant points are, in fact, close to each other. This could, at least for this aspect, makes unnecessary the hypothesis of the inflation.

5) This topology induces verifiable observations:

a) In 2 opposite directions, images of the same galaxy

b) Statistically peaks of galaxies of nearby age of 7×10^9 years. (This proposal seems to be verified: a peak of galaxies around 2200 Mpc seems to exist and allows to date the age of the universe around 14 billion years. However, a verification with a larger statistical series is essential).

c) The NASA/IPAC study of galaxy distances also suggests the verifiable hypothesis that if light travels around in 14 billion years, light capture may sometimes reach us after a second turn. This would then lead to the illusion of seeing galaxies that are more than 14 billion years old and all the more evolved because they are ancient.

d) Galaxies around 7×10^9 years are more similar

e) Galaxies, after going through half the circle tend to converge to a point.

6) The problem of the value of the "critical" mass could disappear and become irrelevant. If there is a centrifugal force, possible because of the 4th dimension of length without contradicting the rule of the geodesic, this force then has effect in

the expansion, or even explains it (To avoid talking about force and remain in the principle of curvature of the GR, it would probably be better to start from entropy). The expansion is not infinite but due to the centrifugal force which, at a given moment, becomes equal to the gravitational force. A balance is made and then fixes a equilibrium density which is equal to what is called the critical density. Overall, in this hypothesis, the rate of expansion decreases with time and will have to be distinguished from the increase in the speed due to gravitation.

7) The force of attraction is always attractive. The shape of this space explains that a repulsive gravitation is only an optical illusion. If in the cosmos, the opposite point is at 7×10^9 LY, as soon as half of this path has been covered, despite the expansion, the masses get closer, then gravity causes a measurable acceleration, an increase in speeds and therefore an impression of acceleration of expansion (this statement also seems to be verified since an acceleration was measured at billion 6 - 7 LY). The shape of this space is well correlated with observations of the movements of Laniakea's galaxies. Black energy retains other theoretical arguments, but is no longer necessary to explain an acceleration of expansion.

8) The black holes in this space have no mathematical singularity in their center since with this topology whatever the level of contraction, there is always a zone in the center that does not belong to space.

5. Limits of This Theory

1) There is first the psychological problem of representing something that is counter-intuitive: the existence of a central zone that does not belong to space (it is necessary to abstain from representing an infinite Cartesian space).

2) This curved space allows to represent a spherical cosmos in 3 dimensions, but it is not said that there is no other representation of the same type (we do not respect the maximum volume given by Luminet [10]). It is not because a theory respects the principle of simplicity that it is necessarily accurate.

3) It necessarily implies since the infinite Cartesian space is virtual and not real that the cosmos is finite. It is therefore in absolute opposition to multiple spaces, "multiverse" and any infinite entity. If we postulate philosophically an infinite cosmos, then this model must be rejected.

4) It is a universe whose radius fluctuates between a maximum and a minimum and which therefore does not necessarily go to the length of Planck.

5) The description of this space does not answer the problem about the nature of finite or infinite time.

6) How to measure a radius of curvature of the order of 1˚ for 40 million LY which could be a confirmation of this model?

Other theories offer explanations for observing very old and evolved galaxies.

This theory eliminates mathematical singularities but conjectures are still necessary for example to prevent the density from increasing without limit.

Conflicts of Interest

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

References

[1] Penrose, R. (1965) *Physical Review Letters*, **14**, 57-59. https://doi.org/10.1103/PhysRevLett.14.57

[2] Bohr, N. (1991) Physique atomique et connaissance humaine, Ed. Gallimard 2015 réed.

[3] Lamé, G. (1859) Leçons sur les coordonnées curvilignes et leurs diverse application, Mallet-Bachelier.

[4] Tegmark, M. (1998) *Annals of Physics*, **270**, 1-51. https://doi.org/10.1006/aphy.1998.5855

[5] Gauss, C.F. (1825) Œuvres complètes, t.IV, p. 189-193.

[6] Sommerfeld, A. (1923) Atomic Structure and Spectral Lines. Londres, Methuen.

[7] Poincaré, H. 1902 () La science et l'hypothèse, chap. III; Les géométries non euclidiennes, Flammarion 2014 réed.

[8] Poincaré, H. (1895) *Journal de l'école polytechnique*, **1**, 1-121.

[9] Poincaré, H. (1904) *Rendiconti del circolo matematico di Palermo*, **18**, 45-110. https://doi.org/10.1007/BF03014091

[10] Luminet, J.P. (2005) L'univers chiffonné, Gallimard/folio.

[11] Luminet, J.P. (2014) La topologie cosmique p94 in Pour la science hors série n˚83 avril-juin.

[12] NED (NASA/IPAC Extragalactic Database)-4D: 394 Distances to 349 Galaxies with v > 1/8 c Compiled from the literature by Barry F. Madore (Pasadena) and Ian P. Steer (Toronto) Updated to January 1, 2007.

[13] Tully, R.B., Courtois, H., Hoffman, Y. and Pomarède, D. (2014) *Nature*, **513**, 71-73. https://doi.org/10.1038/nature13674

[14] Mignonat, M. (2018) *Journal of Modern Physics*, **9**, 1545-1558. https://doi.org/10.4236/jmp.2018.98095

Appendix

Geometric calculation of the position of any point S or event in this finite 3D space:

A system of curvilinear coordinates that seems relatively simple and allows us to remain in a finite space (and therefore in physics) is to consider that:

Every S point in space is marked by three curvilinear coordinates x, y, z.

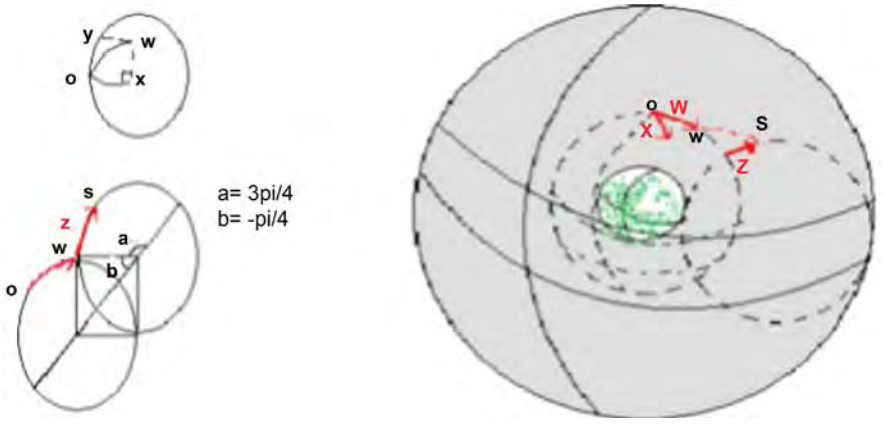
Each curvilinear coordinate x, y, z is located on an arc of circle having the same radius of curvature.

A first 2D sphere is defined by the marks $\boldsymbol{i}$ and $\boldsymbol{j}$. O is a point taken as origin located on this sphere. A first point W is defined by its coordinates x and y. x and y are perpendicular and located on the surface of the sphere. The mark $\boldsymbol{k}$ and so the coordinate z start perpendicular from the point W to the surface of this first sphere.

Any point S of the space can then be defined by a curvilinear coordinate z between -Π/4 and 3Π/4 rd (**Figure 14** left) and a curvilinear coordinate w resulting from the curvilinear coordinates x and y; the coordinate w corresponds to the intermediate point W located on the sphere. By convention WS and OW are located in the same plane and always by convention the curvature for WS is done in the direction dextrogyre when $z > 0$ and in the direction lévogyre when $z < 0$. (The purpose of these conventions is to define a point S by a single group of coordinates and to simplify the calculations. A mathematical generalization would probably be feasible but would not add anything more to the physical representation.)

Starting from OW and applying the first Gauss formula of spherical trigonometry

$(\cos w = \cos x \cdot \cos y + \sin x \cdot \sin y \cdot \cos A)$ we find, since Ox perpendicular to $Oy <=> A = 90^\circ$, that: $\cos w = \cos x \cdot \cos y$



We first do the calculation for z with $0 < z < +3\Pi/4$

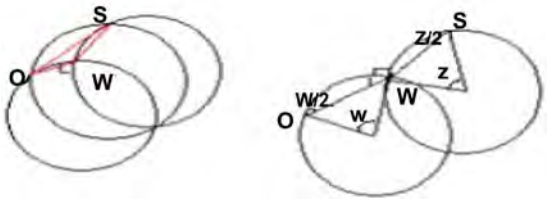
The coordinates w and z are in the same plane and the radius of curvature R is the same for the 2 arcs of circle between OW and WS (drawing above); the radius of curvature R' of OS is different and will be variable according to the

considered point; its calculation is not necessary to position the point S but interesting to do since it gives the curvature of the geodesic which becomes different according to the distance traveled. The angles w, x, y, z are in radians. The lengths of the strings of the circle arcs are:

$$OW = 2R\sin w/2 \tag{1}$$

$$WS = 2R\sin z/2 \quad \text{arc}\,WS \perp \text{arc}\,WO, \text{ so } \quad \text{ang.}OWS = w/2 + \Pi/2 + z/2 \tag{2}$$

$$OS = 2R'\sin s/2 \tag{3}$$



We apply the theorem of Al-Kashi:

$$OS^2 = OW^2 + WS^2 - 2OW \cdot WS \cdot \cos\left(w/2 + z/2 + \Pi/2\right)$$

$$OS^2 = 4R^2\left[\sin^2 w/2 + \sin^2 z/2 - 2\sin w/2 \sin z/2 \cos\left(w/2 + z/2 + pi/2\right)\right]$$

$$OS^2 = 4R^2\left[\sin^2 w/2 + \sin^2 z/2 + 2\sin w/2 \sin z/2 \sin\left(w/2 + z/2\right)\right] \tag{4}$$

$$OS^2 = 4R^2\left[\sin^2 w/2 + \sin^2 z/2 + \sin^2 w/2 \sin z + \sin w \sin^2 z/2\right]$$

$$OS^2 = 4R^2\left[\left(1+\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right] \tag{5}$$

This formula gives the distance of the OS straight segment that can be used in the geometric representation when this finite space is immersed in the infinite Cartesian space that we consider virtual. We keep R which, as part of the expansion of the universe can be replaced by $R = R_0(t)$

(c = speed of light, t = time) We could then write for the coordinates of an event:

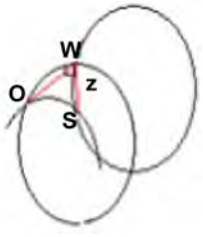
$$OS^2 = c^2t^2 - 4R_0^2(t)\left[\left(1+\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right] \tag{6}$$

If we want to determine the curvature of the geodesic connecting O and S, we can also write:

$$4R'^2 \sin^2 s/2 = 4R^2\left[\left(1+\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right]$$

so:

$$R'^2 = R^2\left[\left(1+\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right]\Big/\sin^2 s/2 \tag{7}$$



Case where $-\Pi/4 < z < 0$

Then the angle $OWS = z/2 + \Pi/2 - w/2$ so,

$$OS^2 = 4R^2\left[\sin^2 w/2 + \sin^2 z/2 + 2\sin w/2 \sin z/2 \sin\left(-w/2 + z/2\right)\right] \quad (8)$$

$$OS^2 = 4R^2\left[\sin^2 w/2 + \sin^2 z/2 - \sin^2 w/2 \sin z + \sin w \sin^2 z/2\right]$$

$$OS^2 = 4R^2\left[\left(1-\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right] \quad (9)$$

The formula giving the right segment OS as a function of time becomes:

$$OS^2 = c^2t^2 - 4R_0^2(t)\left[\left(1-\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right] \quad (10)$$

The formula (7) becomes:

$$R'^2 = R^2\left[\left(1-\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right]/\sin^2 s/2 \quad (11)$$

In total:

Each point of this finished and representative space, multiconnex, with a spherical type and constant curvature and 3 dimensions, can be located by 3 coordinates curvilinear *x*, *y*, *z*:

Knowing that $\cos w = \cos x \cdot \cos y$, then,

When: $0 < z < +3\Pi/4$:

$$\sin^2 s/2 = \left(R/R'\right)^2\left[\left(1+\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right] \text{ or}$$

$$OS^2 = 4R^2\left[\left(1+\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right]$$

When: $-\Pi/4 < z < 0$:

$$\sin^2 s/2 = \left(R/R'\right)^2\left[\left(1-\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right] \text{ or}$$

$$OS^2 = 4R^2\left[\left(1-\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right]$$

If we take a temporal dimension by placing the radius $R = R_0(t)$, we create a space of Minkowski type but which is finite and representative.

An event can be located by the segment OS such as:

$$OS^2 = c^2t^2 - 4R_0^2(t)\left[\left(1+\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right]$$

when $0 < z < +3\Pi/4$

$$OS^2 = c^2t^2 - 4R_0^2(t)\left[\left(1-\sin z\right)\sin^2 w/2 + \left(1+\sin w\right)\sin^2 z/2\right]$$

when $-\Pi/4 < z < 0$

Journal of Modern Physics, 2020, 11, 996-1004
https://www.scirp.org/journal/jmp
ISSN Online: 2153–120X
ISSN Print: 2153–1196

Friedmann-Like Cosmological Equations for the Accelerated Expansion of the Universe

G. Alagar Ramanujam

Former Professor and Head of Physics Department, N.G.M College, Pollachi, India
Email: gravity2003@gmail.com

How to cite this paper: Ramanujam, G.A. (2020) Friedmann-Like Cosmological Equations for the Accelerated Expansion of the Universe. *Journal of Modern Physics*, **11**, 996-1004.
https://doi.org/10.4236/jmp.2020.117062

Received: May 29, 2020
Accepted: July 5, 2020
Published: July 8, 2020

Abstract

In recent papers [1] [2] [3], we framed suitable axioms for Space called Super Space by Wheeler [4]. Using our axioms in Newtonian formalism and considering the density of the universe to be constant in time, we showed in the above references that at $t = 0$ the radius of the universe need not be zero. And thus, we avoided the problem of singularity. We further showed that the Hubble factor is no longer constant in time and goes on decreasing as confirmed by experiments. We pointed out in the above references that Space is the source of dark energy which is responsible for the accelerated expansion of the universe. With a view to improving the above-mentioned results quantitatively, in this paper, we are discussing the consequences of our axioms using Einstein's field equations of general theory of relativity. Friedmann-like Cosmological equations with Dark Energy built-in are derived. This derivation is obtained using Robertson-Walker line element and by introducing a suitable expression for Energy-Momentum tensor in terms of matter and Dark energy contents of the universe. The solutions of our cosmological equations obtained here, show that the radius of the universe cannot reach zero but has a minimum value and there is also maximum value for the radius of the universe. The inflationary expansion of the very early universe emerges from our theory.

Keywords

Axioms for Space, Dark Energy, Seed of the Universe, Einstein's Field Equations, Space (Vacuum) Energy, Accelerated Expansion

1. Introduction

As instruments and techniques improve in sensitivity and accuracy, new observations emerge in science and they challenge the existing concepts and theories and motivate us to find better and more accurate explanations. The experimental

discovery of the accelerated expansion of the universe by Adam G. Riess *et al.*, [5] is a case in hand. Such an accelerated expansion of the universe is clearly beyond the scope of the Big Bang theory and hence provides a strong motivation to go in for a new cosmological theory where acceleration follows as a natural consequence of the theory. The acceleration suggests the presence of an all-pervading anti-gravity force, now called in the literature Dark Energy. In the existing Big Bang theory, the source of Dark Energy which produces the acceleration is completely unknown and this situation is a severe limitation of the theory. In our work reported here, we identify Dark energy as waves in Space whose cause we have extensively discussed. Big Bang theory does not explain how the "cosmic egg" which exploded came into existence. In sharp contrast to this, we have shown that the universe expands from a seed of finite parameters and we have also explained how the seed came into existence. We give below the various opinions of other research workers about the cause and the functional aspect of dark energy and the accelerated expansion of the universe.

In a recent paper, Paolo Christillin, [6] considers the universe as an expanding black hole. To quote Christillin "Backed up also by the consideration of the black body self-energy, the post Big Bang temperature decrease is consistent with particle creation if energy conservation applies at every scale. This process is shown to provide a gravitational repulsive force which can counter balance gravitational attraction, thus allowing the possibility of a steady non-inflationary expansion. That seems to provide an alternative coherent scheme for our picture of the universe evolution, disposing of the cosmological term, of dark energy and of the bulk of dark matter".

According to Gregory Ryskin [7], we have: "Observational data suggest that the universe is spatially flat. Its content dominated by Dark energy which makes up about 73% of the total. The rest is the non-relativistic matter (Including Dark matter); the energy density of relativistic matter and radiation is negligible, except in the early universe. With dark energy is associated negative pressure equal and opposite to its density: this negative pressure acts as a source of anti-gravity, with the result that expansion of the universe is accelerating. The nature of Dark energy remains unexplained".

According to Wheeler [4]: "As we start, we recognise that the arena of Einstein's geometrodynamics is not space, not even spacetime, but super space. Space, to be sure, is the dynamical object. And a spacetime is one classical history of that dynamical object. But the arena in which the dynamics unrolls is super space".

It is highly significant that Wheeler considers the super space as a source of particles. To quote him: "First, we will accept Einstein's general relativity or 'geometrodynamics' in its standard 1915 form, translated of-course the appropriate quantum version. Second, we accept as tentative working hypothesis the picture of Clifford and Einstein that particles originate from geometry; that there is no such thing as a particle immersed in geometry, but only a particle built out

of geometry.

The above paragraphs show that there are diverse views on Dark energy and the link between the particles and the Space. Our views on the link between particles and Space are expressed in our axioms given below for Space.

1) Space is all-pervading and is endowed with potential energy. It has the property of constant self-compression and continually exerting compressive pressure on every system in it.

2) Self-compression results in the formation of infinitesimal spinning quanta of Space, called "formative dust". Due to the surrounding compressive pressure of Space, dust are forced into formation of discrete groups we know as fundamental particles; every group of dust formed by the surrounding compressive pressure of Space has a spin and hence becomes a source of a radial field with a repulsive force at every space-time point.

Explaining our axioms, we state the following: The self-compression of Space results in the infinitesimal spinning quanta of space called dust. A group of dust becomes a fundamental particle with spin. From a spinning fundamental particle, energetic constituent dust are thrown out. Each such thrown out dust produces an outgoing wave in Space. These waves produced by every fundamental particle, constitute the Dark energy. As gravity, there is no place where Dark Energy is not there. The role of Dark Energy is anti-gravity.

The concept of self-compression of Space has a similarity with the concept of action of gravitational field upon itself introduced by Merrison [8]. The above axioms give us a possible methodology for the formation of primordial seed which by its expansion has become the present universe. As already stated, the self-compression of space results in the formation of innumerable dust. Out of the dust emerge fundamental particles, each particle being a group of certain number of dust. Further, certain number of fundamental particles is grouped by the compressive pressure of Space to form a structure. With more and more particles pushed into the structure by the Space, the formation of the structure will continue until the density of the particles in the structure attains homogeneity and isotropy. Inside the Structure the Dark Energy wave from one particle will be pushing every other particle away. Once this anti-gravity force gains an upper hand over the compressive thrust on the structure due to Space, the structure as a whole start expanding. Let β be the radius of the structure when it started expanding. This structure with finite size and density, we consider as the primordial seed of the universe.

When we talk about the primordial seed of the universe the following concept of Vladimir S. Netchitailo [9] is worth considering: "Before the beginning of the World, there was nothing but an External Universe. About 14.2 billion years ago the World was started by a fluctuation in the External Universe and the Nucleus of the World which is a four-dimensional ball was born. An extrapolated nucleus radius at the beginning was equal to basic unit of size, a".

2. Friedmann-Like Cosmological Equations

From our axioms, it is clear that there are two forces operating at the cosmic level. One is the compressive thrust on the universe due to the Space surrounding the universe (Or in other words a kind of gravitational collapse due to gravity). This compressive thrust is a measure of the mass-energy content (W) of the universe. The other force is due to the waves in Space (Dark energy) trying to push the boundary of the universe away. Let E_D be the total Dark energy content of the universe. The dynamics of the universe is controlled by both the mass-energy of the universe (W) and dark energy content (E_D) of the universe. Both W and E_D are functions of time. It must be held in our mind that there is a continuous flow of spinning particles from Space into the Universe. Since every incoming particle carries a mass and becomes a source of waves in Space, W and E_D keep changing. In this context the following observation of Sean M. Carroll [10] of University of Chicago is worth noting: "In a universe with both matter and vacuum energy there is a competition between the tendency of the vacuum energy to cause acceleration and the tendency of matter to cause deceleration, with the ultimate fate of the universe depending on the precise amounts of each component."

Combining Einstein's field equation, R-W line element and the Energy – Momentum tensor $T^{\alpha\beta}$ given as,

$$T^{\alpha\beta} = (\rho + p)u^{\alpha}u^{\beta} - pg^{\alpha\beta}$$

Friedmann [11] obtained his famous cosmological equations which predicted the expansion of the universe. We obtain here similar cosmological equations which take into account the role of Dark energy.

By Einstein's field equations, we have,

$$R_{\nu}^{\mu} - (1/2)Rg_{\nu}^{\mu} = T_{\nu}^{\mu} \tag{1}$$

Our expression for the Energy – momentum tensor is of the form,

$$\begin{aligned} T_{\nu}^{\mu} = &-\left\{\sigma\rho_e / E_D - \alpha / W\rho_m + \left(E_D + W - E_S\right)/3H\right\}U^{\mu}U_{\nu} \\ &+ g_{\nu}^{\mu}\left(E_D + W - E_S\right)/3H \end{aligned} \tag{2}$$

The R-W metric is of the form,

$$\mathrm{d}s^2 = \mathrm{d}t^2 - S^2(t)\left[\mathrm{d}r^2/\left(1-kr^2\right) + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2\right] \tag{3}$$

Here H denotes the Hubble Factor and E_S represents the energy manifested from Space to form the universe. U^{μ} is the four velocity of a galaxy with the conditions, $U^{\mu}U_{\nu}$ equals 1 when $\mu = \nu = 0$ and 0 otherwise. Here α and σ are factors independent of time. ρ_m (mass-energy density), ρ_e (Dark energy density) are considered as constants. $W = (4\pi/3)s^3\rho_m$ and $E_D = (4\pi/3)s^3\rho_e$. We take the unit c (velocity of light) = 1.

It can be verified that the above expression for T_{ν}^{μ} satisfies the divergence condition

$$T_{\nu;\mu}^{\mu} = 0 \tag{4}$$

implying,

$$E_S = \left(W + E_D\right) \tag{5}$$

The above equation is the grand conservation law of Energy between Space and the universe.

In this regard the following observation of Auguste Meesen [12] is worth our consideration: "Since our universe was in a zero-energy state, it is reasonable to assume it arose from vacuum fluctuations. The resulting state was stabilized, but gave then rise to an amazing sequence of transformations and a highly astonishing evolution at different levels of complexification. It began with the conversion of the primeval photon into many material elementary particles, according to the conservation law for u-quantum numbers. These particles interacted with one another, which lead to the formation of compound particles in negative energy state. The total amount of negative energy increased as well as the total amount of positive energy, but energy conservation implies that our universe is still in a zero-energy state." One can find a very mild comparison between the concept of vacuum fluctuations of Auguste Meesen and our concept of particles emerging from Space.

From Equation (1), Equation (2) and Equation (5) we obtain the Friedmann-like cosmological equations as

$$3\left(\dot{s}^2 + k\right)\Big/S^2 = \left(-\sigma\rho_e/E_D + \alpha/W\rho_m\right) \tag{6}$$

$$2\ddot{s}/S + \left(\dot{s}^2 + k\right)\Big/S^2 = 0 \tag{6a}$$

If the R.H.S of the above Equation (6) is positive, then the acceleration will be negative. With negative acceleration the seed cannot grow. For the seed to grow the acceleration has to be positive. So, for the formation of the universe the R.H.S of the Equation (6) has to be negative.

During the early universe, the average spin of the fundamental particles (the source of Dark energy) is very high and therefore a negative pressure due to dark energy remains extremely high when compared to the inward thrust on the universe due to the Space surrounding the universe. Hence the negative term ($-\sigma\rho_e/E_D$) associated with dark energy would have been a dominating factor for the early universe. Hence, retaining the dominating term only, we write for the early universe

$$3\left(\dot{s}^2 + k\right)\Big/S^2 = -\sigma\rho_e/E_D$$

Taking $k = -1$ in the RW metric element, we obtain

$$\dot{s}^2 = \left(1 - \beta/S\right) \quad \text{Where} \quad \beta = \sigma/4\pi \tag{7}$$

Equation (7) shows that for $S < \beta$, $\dot{s}^2$ becomes negative and hence $\dot{s}$ becomes imaginary. The physical meaning of this situation is that expansion starts from $S = \beta$

Solving the above relations, we get

$$S = \left(\beta/2\right)\left(1 - \cosh\phi\right), \text{ with } S > \beta \tag{8}$$

$$t = (\beta/2)(\phi - \sinh\phi) \tag{9}$$

when $S = \beta$, $\phi = \pi$, $\dot{s} = 0$ and $t = (\beta\pi)/2$.

Equation (7) can be written as,

$$H^2 = \left(1/s^2 - \beta/s^3\right) \tag{10}$$

It can be easily verified that H^2 will have maximum value when

$$S = 3\beta/2 \tag{11}$$

Replacing S by $3\beta/2$ in the Equation (10), we get

$$H_m^2 = 4/\left(27\beta^2\right) \tag{12}$$

where H_m is the maximum value of H

When the expansion started, H^2 goes on increasing, reaching a maximum and then decreasing. This shows that the initial expansion was inflationary. Thus, the inflationary expansion of the early universe follows from our theory. It is highly gratifying to note that the inflationary expansion of the early universe and the accelerating expansion of the present universe come out as a consequence of our theory

From the Equation (6a) and Equation (6), we get

$$\ddot{s} = \beta/2s^2 \tag{13}$$

The acceleration $\ddot{s}$ goes on decreasing with distance but remains always positive. This decrease is understandable. As time evolves, the average spin of the fundamental particles in the universe goes on decreasing and therefore the power of the dark energy waves produced by the fundamental particles decreases leading to less and less acceleration. Gradually as the power of the dark energy decreases, the universe enters Space or Matter dominating era, where the compressive thrust on the universe exceeds the repulsive thrust due to the dark energy. Let S_r be the radius of the universe when it enters the matter dominating era. In this case, the nature of curvature of Space changes and k becomes zero making the acceleration negative.

For this era, we have,

$$3\dot{s}^2/S^2 = \alpha/W\rho_m \quad \text{with} \quad S > S_r$$

$$\dot{s}^2 = \alpha S^2/(3W\rho_m)$$

$$\dot{s}^2 = \gamma/S \quad \text{where} \quad \gamma = \alpha/4\pi\rho_m^2 \tag{14}$$

The above Equation (14) can be solved to get,

$$S = \left(2/3\sqrt{\gamma}\right)^{2/3} t^{2/3} \quad \text{and}$$

$$\ddot{s} = -(2/9)\left(2/3\sqrt{\gamma}\right)^{2/3} t^{-4/3}$$

Thus, when $S > S_r$, k becomes zero and acceleration becomes negative. As time further increases, the compressive thrust on the universe due to Space becomes enormously high. Consequently, at a point of time the nature of curvature inside

the universe changes in such a way that k becomes +1. For the era when $k = +1$ we have,

$$3\left(\dot{s}^2+1\right)\Big/S^2=\alpha\left(W\rho_m\right) \quad \text{with} \quad S\gg S_r \tag{15}$$

$$\mathrm{d}s^2=\mathrm{d}t^2-S^2\left(t\right)\left[\mathrm{d}r^2\Big/\left(1-r^2\right)+r^2\mathrm{d}\theta^2+r^2\sin^2\theta\mathrm{d}\phi^2\right]$$

Now the Friedmann-like cosmological equation becomes,

$$\dot{s}^2=\left(\gamma/S\right)-1 \quad \text{where} \quad \gamma=\alpha\Big/\left(4\pi\rho_m^2\right) \tag{16}$$

The solution for the above equation during the Space dominating era can be written as

$$S=\left(\gamma/2\right)\left(1-\cos\theta\right) \quad \text{with} \quad S>S_r \quad \text{and}$$

$$t=\left(\gamma/2\right)\left(\theta-\sin\theta\right)$$

During the matter dominating era, the acceleration becomes negative, and hence the expansion of the universe will have deceleration and eventually will stop expanding at a later time. The dynamical behaviour of the universe after its expansion stops at the distance $S = \gamma$, is under investigation.

3. Conclusions

The cosmological model given in this paper, which we call "Space Transformation Model", is based on our two axioms for Space. These two axioms are condensed versions of the views of Vethathiri Maharishi on Space [13].

The primordial seed which expanded into the present universe comes from and of the Space. Due to the dominating dark energy, the seed started expanding when $S = \beta$ with positive acceleration. The acceleration remains positive but went on decreasing. When the radius of the universe reaches the value S_r, the velocity of the expansion will be $(1 - \beta/S_r)^{1/2}$. Beyond S_r, the compressive thrust on the universe due to space will dominate the repulsive force due to the dark energy and hence the curvature of Space inside the universe keeps changing so that from −1, k changes to 0 and then to +1. In both cases when $k = 0$ and +1, the acceleration of the expansion becomes negative. Hence, the velocity of the expansion will start decreasing, finally becoming zero when $S = \gamma$. In the Big Bang model, during the evolution of the universe, the factor k in R-W metric remains either −1 or 0 or +1 throughout the evolution of the universe. But in our theory presented here, the universe during its expansion assumes all the three phases $k = -1$, 0 and 1 one after the other. Once expansion velocity becomes zero at $S = \gamma$, a natural question will be: "what will be the dynamics of the universe after that?" Our theoretical investigation regarding this question will be published elsewhere in future.

Inflationary expansion of the early universe when it started expanding and the acceleration of the expansion of the universe at present turn out as the consequences of our theory.

The Hubble constant which was once considered as a universal constant, as

recent observations show, is no longer a constant. We have obtained here a relation between H and S in Equation (10) of this paper which gives the nature of variation of Hubble factor with radius of the universe. Such a variation of the Hubble constant has been obtained also by Mishra and Vadrevu [14].

By introducing the term Es in our Energy-Momentum tensor, we established that Space is the source of both the matter and the dark energy contents of the universe.

Acknowledgements

I express my deep sense of reverence to Shri. Vethathiri Maharishi, for a number of personal discussions on the nature of Space and gravity. My sincere thanks are due to Arulnidhi Uma Vethathiri, Arulnidhi Mr. Vijay Arora (USA) and Arulnidhi Mr. Raj Taneja (USA) for the very useful discussions regarding the focal theme of this paper. It is a pleasure to thank K. Anbarasan, D. Padma Priya and Dr. Shanmuga Priya of Chennai for the very many useful discussions.

Conflicts of Interest

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

References

[1] Alagar Ramanujam, G., Fitzcharles, K. and Muralidharan (2017) *Journal of Modern Physics*, **8**, 1067-1071. https://doi.org/10.4236/jmp.2017.87068

[2] Alagar Ramanujam, G., Fitzcharles, K. and Muralidharan (2019) *Indian Journal of Physics*, **93**, 959-963. https://doi.org/10.1007/s12648-018-01364-9

[3] Arora, V., Taneja, R. and Alagar Ramanujam, G. (2019) *International Journal of Trend in Research and Development*, **6**, 472. http://www.ijtrd.com/papers/IJTRD20708.pdf

[4] Wheeler, J.A. (1969) Relativity. *Proceedings of the Relativity Conference in the Midwest*, Cincinnati, 2-6 June 1969, 31-42. https://www.springer.com/gp/book/9781468407235

[5] Riess, *et al.* (1998) *The Astronomical Journal*, **116**, 1009-1038. https://iopscience.iop.org/issue/1538-3881/116/3 https://doi.org/10.1086/300499

[6] Christillin, P. (2016) *Journal of Modern Physics*, **7**, 1331-1344. https://doi.org/10.4236/jmp.2016.711119

[7] Ryskin, G. (2015) *AstroParticle Physics*, **62**, 258-268. https://doi.org/10.1016/j.astropartphys.2014.10.003

[8] Merrison, J.P. (2016) *International Journal of Astronomy and Astrophysics*, **6**, 312. https://doi.org/10.4236/ijaa.2016.63026

[9] Netchitailo, V. (2016) *Journal of High Energy Physics, Gravitation and Cosmology*, **2**, 593-632. https://doi.org/10.4236/jhepgc.2016.24052

[10] Carroll, S.M. (2001) *Living Reviews in Relativity*, **4**, 1. https://doi.org/10.12942/lrr-2001-1

[11] Friedmann (1922) *Zeitschrift für Physik*, **10**, 377-386. (In German)

https://doi.org/10.1007/BF01332580

[12] Meesen, A. (2017) *Journal of Modern Physics*, **8**, 35-58.

[13] Vethathiri, Y. (1995) Unified Force. Vethathtiri Publications, Erode.

[14] Mishra, B. and Vadrevu, S. (2017) *Astrophysics and Space Science*, **362**, 26. https://doi.org/10.1007/s10509-017-3006-2

Journal of Modern Physics, 2020, 11, 1005-1012
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Some Equations for the External Space and Their Implications on the Inverse Cone Concept and the Geometry of the Physical Universe

J. G. Lartigue

Retired Professor, National University of Mexico, Mexico City, Mexico
Email: jmlg5@hotmail.com

How to cite this paper: Lartigue, J.G. (2020) Some Equations for the External Space and Their Implications on the Inverse Cone Concept and the Geometry of the Physical Universe. *Journal of Modern Physics*, **11**, 1005-1012.
https://doi.org/10.4236/jmp.2020.117063

Received: May 6, 2020
Accepted: July 7, 2020
Published: July 10, 2020

Abstract

In previous works, they were proposed a photonic model of the Big Bang [1] and several parameters derived from the Hubble-Lemaitre equation [2]. Since these parameters result higher than the classical ones and, otherwise, the General Theory of Relativity does not apply far away the Physical Universe, in this paper, it will be revised the adequacy of such parameters in the external Space and their influence on the relativistic concept of the cone of time. As well, it will be intended to define the Physical Universe geometry accordingly to a thermo-dynamical analysis of the Big Bang.

Keywords

De Sitter Equation, The Inverse Time Cone, Big Bang Characteristics

1. Background

From to the Schwarzchild rule, the Luminosity of the Physical Universe should have been $\mathfrak{L}_{pu} \sim 1.8 \times 10^{59}$ (erg/s) which, multiplied by the Big Bang duration, 10^{12} (s) [3], gives an energy production $\sim 2 \times 10^{71}$ (erg). However, the energy of to the original CMB radiation would be much higher: the present density of the CMB relic is about 416 (photons/cm^3) [4]; applying the cosmological principle, it is obtained a total number of CMB photons, in the total Universe volume (3×10^{85} cm^3), of 1.2×10^{88} (photons) with a heat content of 3×10^{88} (erg). By assuming that the final temperature of the Big Bang period was $\sim 2 \times 10^4$ (K), that would mean a z factor about 2×10^3 which implies an original energy of the CMB $\sim 6 \times 10^{91}$ (erg). Therefore, the fraction of the total Big Bang energy devoted to the creation of the Physical Universe was $\sim 10^{-20}$ and the Luminosity corresponding to the total Universe age would now be: $\mathfrak{L}_u \sim 1.3 \times 10^{74}$ (erg/s).

The Einstein's equation of General Relativity was modified by De Sitter for a powder cloud as:

$$\boldsymbol{R}^{ij} - \left(\frac{1}{2}\right) g^{ij}\boldsymbol{R} = 8\pi G\rho \boldsymbol{v}^{ij} + \Lambda g^{ij} \quad (1).$$

Since $\rho = 0$ in the external Space and Λ was discarded by Einstein, the volumetric accelerations terms would cancel out. Otherwise, though the gravitational intensity does not vanish at distances even higher than $\boldsymbol{r}_o$, the expansion velocity of the vacuum Space is not a function of the $\boldsymbol{g}$ value; it depends on the universal constant acceleration ($\boldsymbol{\Gamma}_H$) and time.

Therefore, it could be useful to summarize some equations that may be applied to describe the external Space dynamics. Those are:

- The Hubble parameter:

$$H = 2/t\left(\mathrm{s}^{-1}\right). \quad (2)$$

- The Hubble velocity:

$$\boldsymbol{v}_H = H \cdot \boldsymbol{r} = 2\boldsymbol{r}/t\left(\mathrm{cm/s}\right). \quad (3)$$

- The constant Hubble acceleration:

$$\boldsymbol{\Gamma}_H = H^2\boldsymbol{r}/2\left(\mathrm{cm/s}^2\right). \quad (4)$$

- The Space expansion velocity:

$$\boldsymbol{v}_s = \boldsymbol{\Gamma}_H \cdot t\left(\mathrm{cm/s}\right). \quad (5)$$

- The Hubble field potential:

$$V_H = H^2\boldsymbol{r}^2/4\left(\mathrm{cm}^2/\mathrm{s}^2\right). \quad (6)$$

Simultaneously, it has been assumed [2] that the Big Bang was a singularity, *i.e.* the most probable origin of time, Space, matter and the physical laws. So, the Space started immediately an expansion, a fact discovered by Hubble-Lemaitre in 1929 [5] and confirmed as an accelerated movement by Riess *et al.* [6]. The author assumed the simplest case: that of a constant acceleration, as a function of distance and time. Besides it was proposed an equation for ɣ to avoid an imaginary proper time if $\boldsymbol{v} > \boldsymbol{c}$, [2]. However, the possibility of an imaginary time has been mentioned by Hawking [7] as a component of the Euclidic Space, adequate to describe the expansion velocity of the Space.

The Poisson equation for the Hubble potential was obtained in reference [1] as

$$\nabla^2 V_H = H^2\left(\mathrm{s}^{-2}\right), \quad (7)$$

which permits to express the scalar of curvature of the free Space as:

$$R = 6H^2\left(\mathrm{s}^{-2}\right). \quad (8)$$

The Einstein Equation (1) as a function of a Hubble tensor is:

$$G_{\mu\nu} = g_{\mu\nu}3H^2. \quad (9)$$

The scalar of this tensor would be:

$$T^H = 3H^2 = \Lambda\,. \tag{10}$$

2. The Variation of the Cone of Time

In a previous work [8] it was concluded that, in the Physical Universe, the light velocity does not depend on the gravitational potential, *i.e.* it is a universal constant. Besides, since the Space expands at growing velocities (including the matter of the Hubble Flux) it is very probable that the light accelerates simultaneously as the Space does; this would imply an apparent dilemma: if the Space expansion work is carried out by only the Hubble potential (V_H), without any change in the photon's energy or, opposed, the photon's work implies a redshift: $W = h{\cdot}\Delta\nu$ that would limit the photon's life-time. In the author's opinion all works are done at expenses of V_H, without any wear of the expanding matter.

The Special Relativity concept of the inverted cone of time is based on the principle of the light speed constancy; so, the radius of the cone should be calculated by the straight line equation: $\boldsymbol{r} = \boldsymbol{c}{\cdot}t$ (10). If the light would travel in the external Space, at the same velocity than Space does, the distance should be determined by the equation:

$$\boldsymbol{r} = \frac{\boldsymbol{\Gamma}_H}{2}\cdot t^2 \tag{11}.$$

Substitution of the acceleration constant drives to the equation:

$$\boldsymbol{r} = 10^{-7}t^2 \tag{12}.$$

i.e. the cone slant follows a parabolic trajectory respect to time, as shown in Figure 1. In the interval from $t = 0$ to $t = 10.5$ Gy, the radius of the straight cone basis was higher than the parabolic one though, after such a time, the parabolic

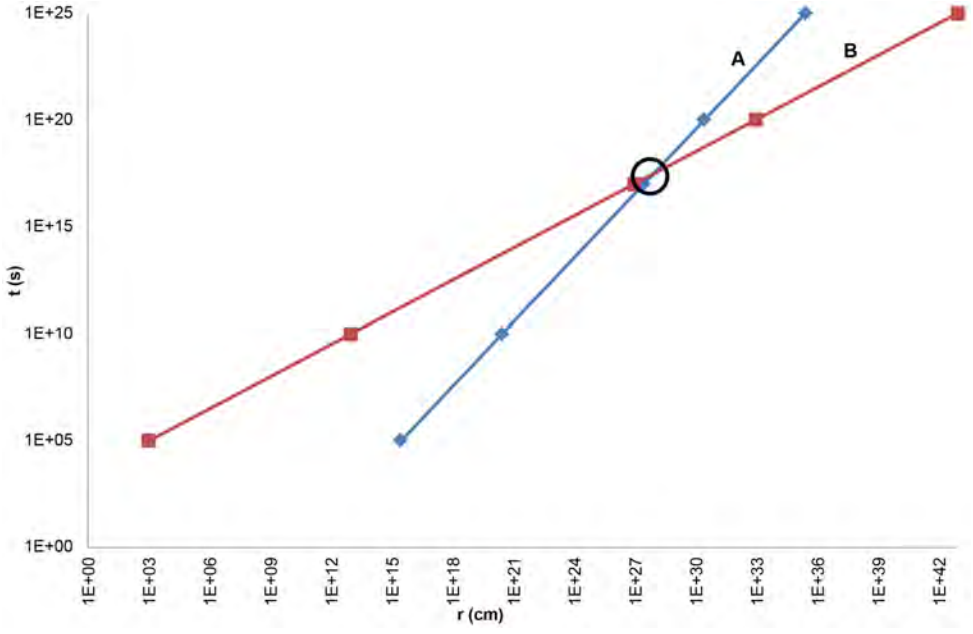


Figure 1. The slants of the straight (A) and the parabolic (B) cones of the Space Expansion. Both lines intersect at an Universe age of ~10^{10} (y).

cone radius grows faster than the straight one. Even so, at any future Space velocity, the so-called "elsewhere zones" will never disappear.

3. The Lorentz Transform in an Accelerating Space

The basic equation of the Lorentz transform is:

$$\boldsymbol{v}'_x = \boldsymbol{v}_x - \boldsymbol{v}; \tag{13}$$

$\boldsymbol{v}$ is the separation velocity of two inertial frames of reference (primed and not-primed) and the sub index x refers to the same coordinate.

The Space accelerated expansion does not correspond to an inertial frame, though in the Physical Universe they may be present both types of frames. So, it could be proposed the following equation:

$$\boldsymbol{v}_x = \boldsymbol{v}'_x + \boldsymbol{\Gamma} \cdot t \tag{14}$$

where: $\boldsymbol{v}_x$ = velocity of matter in the Space,

$\boldsymbol{v}'_x$ = velocity of matter in the Physical Universe

$\boldsymbol{\Gamma} \cdot t$ = expansion velocity of Space at time t. Each one of the application points of this vector corresponds to a co-moving coordinate.

Since Equation (14) is a vector sum, the velocity of matter $\boldsymbol{v}_x$ could be lower or higher than that of the Space. Given the order of magnitude of such velocities, it is probable that the matter velocity $\boldsymbol{v}'_x$ would always be considered as a peculiar velocity.

4. The Universe Form

Since the Space expansion occurred in all directions around the Big Bang point (same as it did the Physical Universe) and if this one is composed of the known matter, all of that must now be confined into a ring: a spherical shell geometry, as a spherical tokamak. The nice diagrams of the Universe development, such as that of reference [6], show the History of the Universe expansion into an angle of <60˚. If it would be extended to a 4π geometry, it could be confirmed that all of the surviving matter must now be confined into a spherical shell. It means that the condensation of radiation would have started when the Universe temperature must be lower than 2×10^4 (K), *i.e.* at the end of the Big Bang. So, as shown in the Universe cross section of **Figure 2**, if $\boldsymbol{r}_B$ was the Big Bang radius, the two radii ($\boldsymbol{r}_i$ and r_e) correspond to the width of the Physical Universe; then, $\boldsymbol{r}_i$ would be the radius of the past expanding Space. Consequently a complete light turn, lengthways the annular Physical Universe, would take a time $t_{pu} \sim 13 t_o$.

The condensation of photons to leptons could not be possible at the initial Big Bang temperature of 10^{31} (K), *i.e.* much higher than the binding energies of leptons (~1 eV). Therefore such a condensation would have started at the end of the Big Bang period ($t_B \sim$ 32,000 years) [3]. The necessity of photonic condensation opposes to the common theory of a short Big Bang (with small dimensions and a high temperature) which required to be followed by a short inflationary period, to expand the photonic volume in order to low the temperature till the

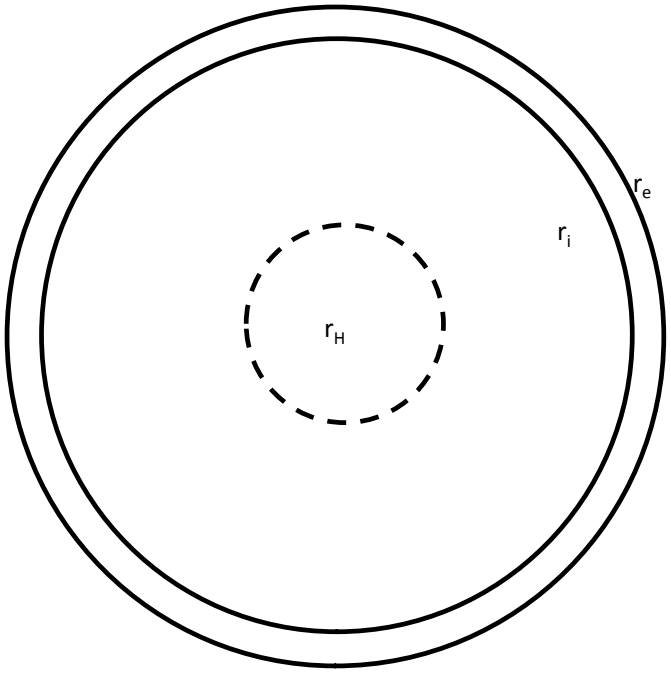


Figure 2. A cross section of the spherical Universe (not at scale); r_B is the radius of the Big Bang ~ 10^{17} (cm); r_e is the estimated radius of the Space~ 2×10^{28} (cm) and r_i is the internal radius of the Physical Universe. The Observable Universe thickness is $(r_e - r_i) \sim 2 \times 10^{19}$ (cm).

necessary level for matter condensation. Instead, at the end of the long Big Bang period proposed by reference [3], the photon gas temperature must have been in the order of 2×10^4 K, adequate to permit the lepton's condensation. So, the Inflationary period should not be necessary [9].

5. Some Consequences of the Big Bang

In a previous work [1] it was assumed the Universe expansion as adiabatic since there is not any enter or loss of heat in the total Universe. It is important to point out several aspects already mentioned. The radial expansion velocity of the Big Bang is given by Equation (5); so, at the end of the Big Bang lifetime, such velocity was only ~ 2×10^5 (cm/s). Anyway, the constancy of $\boldsymbol{c}$ was kept by every one of the original photons across their continuous internal collisions. The Big Bang final radius would be $\boldsymbol{r}_B \sim 10^{17}$ (cm) and the energy density would correspond to 5×10^{45} (erg/cm^3). Consequently, the mean free path at the end of the Big Bang must have been too small to let the existence of a high fraction of low energy photons; these ones would increase and react in the subsequent period ($t_c - t_B$). Besides, accordingly to reference [10], the ratio of the present radial function to that of the Big Bang end implies a z value ~ 5×10^3.

Obviously it must have existed a mixture of frequencies and energies in the original photons which, later, devoted a small fraction $(\sim 10^{-20})^*$ to condense a Physical Universe and, simultaneously, to loose energy by a redshift process and by collisions with the condensed matter, so producing the today known CMB

spectrum [11]. Besides, the energy devoted to the Physical Universe creation could have generated the necessary number of Higgs photons (125 GeV/photon ~ 0.2 erg/photon) which corresponds to a temperature ~ 10^{15} (K), *i.e.*, well inside the Big Bang.

In reference [1] it was applied the convective operator to the photon's energy to obtain the equation of continuity, which gives a luminosity, at the t_c time, of L ~ 4×10^{80} (erg/s).

- The energy equivalent of the assumed mass (3×10^{58} g) of the Physical Universe is about 2.2×10^{77} (erg).

6. Some Necessary Metaphysical Speculations

To date, the existence of the cosmic radiation background has been assumed as the best proof of a Big Bang origin of the Universe. Such conclusion has been adopted in this and a previous paper [1] that included, necessarily, several metaphysical concepts to try to explain some aspects of the creational task. So, the *creation-ex-nihilo* started with the Big Bang, whose characteristics have been summarized above, including the creation of time, space, matter and physical laws. Based on thermo-dynamical concepts of a photonic plasma, some steps have been mentioned as functions of temperature. Assuming that the Creator (or the creative force) had imposed himself an ethical condition, this one was transmitted to some of the created beings: those supplied with a conscience. So, the transcendence characteristic of the Creator was impressed in the eternal soul of the human beings; and his ethical quality was transmitted by means of the conscience and the free will. However, though the conscience would correspond to the soul, its ethical use depends on the human discernment, i. e. an intellectual capacity. The search for knowledge seems to be another characteristic of the human conscience (as a probable consequence of the original respect for the Tree of Life), a task without any time limit.

7. Conclusions

1) The relationship between the Hubble parameter and time has been given by Equation (5) as $H = \sqrt{2}/t\left(\mathrm{s}^{-1}\right)$.

2) The Space expanding acceleration has been expressed as a constant by equation (4) as $\Gamma_H = 2\times 10^{-7}\left(\mathrm{cm}/\mathrm{s}^2\right)$. **Figure 1** shows the straight and the parabolic slants of the cones of time, which intersected at t ~ 1.5×10^{10} (y). So, the parabolic cone radius is today bigger than the previously assumed cone.

3) The positive Hubble potential of the Space expansion was deduced in Equation (7): $V_H = H^2 r^2/4$ as a growing function of the scale factor.

4) The time when the Space expansion reached the $\boldsymbol{c}$ velocity, t_c, has been found to occur at 1/3 of the present time t_o. Obviously, in the subsequent period (t_o – t_c), the total Space expanded at higher velocities than $\boldsymbol{c}$, carrying together the Physical Universe; though, inside this one, the $\boldsymbol{c}$ value of matter remained as a constant. Besides, after the t_c time, it had been deduced the possibility of an

imaginary time domain [7] that would be the characteristic coordinate in an Euclidic Space [8]. Assuming that the present radial function of the Physical Universe would correspond to the most distant object detected, with $z \sim 12$; and also assuming that after the t_c time the most external objects are traveling at higher velocities than $\boldsymbol{c}$, it could be necessary to apply similar equations to the above named (2 - 14). Otherwise, the confirmation of the Physical Universe expansion must have probably been based on an unexpected speed of the ($z \sim 12$) objects; so this would be, to date, the limit of the Physical Universe. Assuming that such observation was made from the middle of the spherical shell, it means a thickness of this shell of $\boldsymbol{r}_{pu} < 2 \times 10^{19}$ (cm) as shown in **Figure 2** for the Physical or Visible Universe.

5) The condensation of photons in leptons could only have occurred at the end of the Big Bang duration (~ 32,000 y) when the temperature was lower than 2×10^4 (K), *i.e.* the binding energy of leptons. After, they continued the synthesis processes of quarks, atoms and stars as well as manifestations of the different types of forces, to integrate the present Physical Universe. These conclusions imply that the generally assumed inflationary period, following a short Big Bang, should not be necessary. Furthermore, for simplicity, it has been supposed that the creation of original photons stopped at the Big Bang end. So, because of the **c** constancy inside the Physical Universe, this has acquired a ring geometry (**Figure 2**): so, a small fraction of photons started a mass condensation for the Physical Universe though, most of them, suffered a redshift process till the present CMB spectrum. Therefore, the mass generation would have occurred between the ($t_c - t_B$) period, by nuclear reactions and later, by electro-magnetic and gravitational forces.

6) Related to metaphysical propositions (already discussed), it has been concluded that the *creation ex-nihlo* started with the Big Bang, together the time, Space, matter and physical laws. The first law could have been the expansive acceleration of Space and, consequently, that of matter. Besides, the transcendence and ethical characteristics of the Creator were transmitted to the human kind by means, respectively, of an eternal soul and a conscience (with a free will). The eternal search for knowledge seems also to have been impressed in the human conscience.

Conflicts of Interest

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

References

[1] Lartigue, J.G. (2018) *Journal of Modern Physics*, **9**, 2443-2456.

[2] Lartigue, J.G. (2019) *Applied Physics Research*, **11**, 70. https://doi.org/10.5539/apr.v11n4p70

[3] Hogan, C.J. (1999) Energy Flow in the Universe.

[4] Leff, H.S. (2002) *American Journal of Physics*, **70**, 795.

https://doi.org/10.1119/1.1479743

[5] International Astronomical Union (2018) Resolution B4.

[6] Riess, A., *et al.* (1998) *Astronomical Journal*, **116**, 1009-1038. https://doi.org/10.1086/300499

[7] Lartigue, J.G. (2020) *Applied Physics Research*. (In Press)

[8] Hawking, S. (1988) Historia del Tiempo. Ed. Crítica, México, 179.

[9] Hossenfelder, S. (2017) Is the Inflationary Universe a Scientific Theory? Not Anymore. https://www.forbes.com/sites/startswithabang/2017/#77d4d994370a

[10] Sartory, L. (1996) Understanding Relativity. UCLA, 303.

[11] Sartory, L. (op. cit.) Figure 9.18.

Journal of Modern Physics, 2020, 11, 1013-1030
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Conceptual Deficiency for the Use of Christoffel Connection in Gravity Theories

Georgios Kofinas

Research Group of Geometry, Dynamical Systems and Cosmology, Department of Information and Communication Systems Engineering, University of the Aegean, Karlovassi, Samos, Greece
Email: gkofinas@aegean.gr

How to cite this paper: Kofinas, G. (2020) A Conceptual Deficiency for the Use of Christoffel Connection in Gravity Theories. *Journal of Modern Physics*, **11**, 1013-1030.
https://doi.org/10.4236/jmp.2020.117064

Received: January 29, 2020
Accepted: July 11, 2020
Published: July 14, 2020

Abstract

Christoffel connection (or Levi-Civita affine connection) did not enter gravity as an axiom of minimal length for the free fall of particles (where anyway length action is not defined for massless particles), nor out of economy, but from the weak equivalence principle (gravitational force is equivalent to acceleration according to Einstein) together with the identification of the local inertial frame with the local Lorentz one. This identification implies that the orbits of all particles are given by the geodesics of the Christoffel connection. Here, we show that in the presence of only massless particles (absence of massive particles), the above identification is inconsistent and does not lead to any connection. The proof is based on the existence of projectively equivalent connections and the absence of proper time for null particles. If a connection derived by some kinematical principles for the particles is to be applied in the world, it is better for these principles to be valid in all relevant spacetime rather than different principles to give different connections in different spacetime regions. Therefore, our result stated above may imply a conceptual insufficiency of the use of the Christoffel connection in the early universe where only massless particles are expected to be present (whenever at least some notions, like orbits, are meaningful), and thus of the total use of this connection. If in the early universe, the notion of a massive particle, which appears latter in time, cannot be used, in an analogous way in a causally disconnected high-energy region (maybe deep interior of astrophysical objects or black holes), the same conclusions could be extracted if only massless particles are present.

Keywords

Particle Kinematics, Massless Particles, Equivalence Principle, Local Inertial Frames, Local Lorentz Frames

1. Introduction

Christoffel connection (also known as Levi-Civita connection) is undoubtedly the dominant affine connection in gravitational physics. Einstein's General theory of Relativity is the leading paradigm of gravitational theories and is constructed out of this connection. In recent years, there was a renewed interest and activity in generalizations or modifications of Einstein gravity. In one direction, various functions of the Ricci scalar acted as Lagrangians, instead of merely the Ricci scalar itself ([1] [2] [3] [4] and references therein). Alternative gravitational theories built up from higher-order curvature corrections, other than the Ricci scalar, have also appeared. Namely, using scalar quantities of the Riemann tensor, new theories were constructed [5] [6] [7] [8]. The Gauss-Bonnet term is a particular combination of such curvature invariants and appears in low-energy effective action of string theory [9] [10]; when this term is coupled to a scalar field, non-singular cosmological solutions arise [11] [12] [13]; a function of Gauss-Bonnet term can also serve as Lagrangian to construct viable models [14] [15] [16]. Greater-than-fourth-order theories, containing derivatives of curvature invariants, were another option [17] [18]. Non-relativistic toy model theories of quantum gravity were also proposed [19] [20]. Scalar fields were included in various scalar-tensor theories [21] [22] [23]. In addition, higher-dimensional gravity theories of the Kaluza-Klein form [24] [25] or of the braneworld scenario [26] [27] have attracted much interest over the years. In all the above theories, the standard approach for deriving the field equations is the metric formalism, where the Christoffel connection is used (the Palatini formalism, where the connection is treated as independent variable, is another approach which is also adopted sometimes).

For positive-definite metrics of Riemannian geometry, the geodesics of the Christoffel connection provide the curves of minimal (sometimes maximal) length as is manifested by the variation of the length action $\int \mathrm{d}\sigma \sqrt{g_{\mu\nu}\frac{\mathrm{d}x^{\mu}}{\mathrm{d}\sigma}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\sigma}}$ for some parameter σ along the orbits $x^{\mu}(\sigma)$. Similarly, in spacetime physics, the action $\int \mathrm{d}\sigma \sqrt{-g_{\mu\nu}\frac{\mathrm{d}x^{\mu}}{\mathrm{d}\sigma}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\sigma}}$ extremizes the proper time (or the length in some sense) of timelike curves, which are the curves of massive particles, and provides the geodesics of the Christoffel connection as the corresponding orbits. However, for massless particles the proper time is zero, a photon orbit on the null cone has zero length, so a similar action is not defined. Therefore, the axiom of minimum length for massive particles is accomplished by another axiom, that of continuation of the geodesic equation of massive particles to be also valid in the massless case. This fact does not mean that the geodesic equation for all particles does not arise from an action. It arises e.g. from the action $\int \mathrm{d}\sigma \left(e g_{\mu\nu}\frac{\mathrm{d}x^{\mu}}{\mathrm{d}\sigma}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\sigma} - e^{-1}m^2 \right)$, where m is the particle mass and e a one-dimensional vierbein. This last action is meaning-

ful even in the massless case, where now the nullity constraint $g_{\mu\nu}\frac{\mathrm{d}x^{\mu}}{\mathrm{d}\sigma}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\sigma}=0$ provides through multiplication with the Lagrange multiplier e a meaningful (off-shell) Lagrangian to be varied (actually in agreement with the method of the theory of constraints). However, this last action is still an axiom with purpose to get the geodesic equation of the Christoffel connection as providing the orbits of all particles. In a similar way the minimal surface action of string theory is assumed. The fact that the Christoffel connection is the simplest one, which only contains the metric part and not extra degrees due to torsion or non-metricity, does not necessarily declare it as the physical connection to describe the universe (anyway, our job is not topography to stick to the notion of minimal length). A lot of modifications of General Relativity contain various extra fields in the action without enriching the Christoffel connection with extra fields, therefore the economy argument of the Christoffel connection is immediately canceled. Moreover, people today are more concerned about the particle content of gravity theories built up from the Christoffel or other connections and rarely follow Einstein's spirit that principles on particle kinematics could provide a connection to build up a theory. In a sense, string theory is closer to this spirit since it starts with the motion of the string to construct the theory. At the opposite end is the first order formulation of various gravity theories, where the metric (or the vierbein) and the connection are considered quite independent; the cost however from this approach is that the connection (or equivalently the torsion or non-metricity) carry in general a large number of new components and it is difficult to believe that all of them have some physical meaning. An intermediate situation probably looks more promising.

Christoffel connection entered physics through Einstein, who certainly did not adopt it axiomatically or in a geometric manner, but he had to be convinced for its virtue out of physical arguments. Of course, one can still keep on using the Christoffel connection even without any justification, but our aim here is to try to reduce its physical significance out of theoretical reasons which invalidate Einstein's arguments in some probably realistic spacetime regions. To be more precise, the geodesics of the Christoffel connection as the orbits of all particles arise by the identification of the local inertial frame/freely falling frame (ingeniously discovered by Einstein refining the notion of weak equivalence principle) with the local Lorentz frame/local Minkowski frame (today the existence of such a frame probably does not need any justification since the notion of a spacetime metric which generalizes the Minkowski metric is considered fundamental, but it seems as well that it can consistently be introduced through what is called Einstein equivalence principle concerning the non-gravitational laws of physics).

In this work, we prove that in the presence of only massless particles (massive particles do not exist) the above identification is meaningless and does not lead to any connection. In an alternative construction of a connection, responsible for providing the orbits of all particles, the existence of the above two frames can

usually be adopted, but not their identification. Our proof is based on the existence of projectively equivalent connections and the absence of proper time for massless particles. The fact that the relevant group for massless particles is the conformal group does not affect the validity of our result since the Lorentz group is still present. We argue that this result is probably relevant in our world. It seems that the most direct and solid regime for the applicability of the above result is the early universe. In the early universe, the standard approach is to consider the contributions from loops and thermal corrections at finite temperature. As a result, all particles are expected to be massless before the electroweak symmetry breaking. Even for theories with phase transitions at higher energy scales, before these scales all particles are still expected massless. If these massless particles have to decide about their motion in a gravitational environment out of some kinematical principles (at least as long as semiclassical notions such as orbits are still defined), they cannot be based on the notion of a massive particle which arises later in time. Thus, the Christoffel connection in this regime is not the result of physical arguments on particle kinematics, thus probably losing its overall reliability (in this same spirit, even the axiom of continuation to the purely massless case of the minimal length geodesics of the massive particles appeared later in time, is also insufficient). On the contrary, if another connection could be defined out of physical arguments concerning the kinematics of all particles, which arguments will equally well be meaningful everywhere in spacetime, we consider that this new connection could be a more appropriate and realistic connection to be valid everywhere in spacetime. Moreover, the existence of the notion of orbit (either classical or semiclassical) seems still reasonable around the TeV scale since the full strong quantum gravity regime is normally expected at much higher energies. Another case where the insufficiency of the Christoffel connection may occur is in spacetime regions of very high energy and temperature where only massless particles may exist, which regions are causally disconnected from the rest of spacetime where massive particles are also present (deep interior of supermassive astrophysical objects or even black holes are not impossible to be such regions). The above thoughts could also be relevant in the presence of extra dimensions with some fundamental mass scale.

2. The Local Inertial Frames

The local indistinguishability of gravitational and inertial forces, which is an essential step for the derivation of Einstein's General Relativity (and therefore of other theories as well), is described by the axiom that there exists a coordinate system χ^μ around the event p (local inertial frame or freely falling frame) and a parameter λ, such that the orbits $\chi^\mu(\lambda)$ of freely falling test particles obey

$$\frac{\mathrm{d}^2\chi^\mu}{\mathrm{d}\lambda^2}(p)=0. \tag{2.1}$$

The orbit $\chi^\mu(\lambda)$ has locally the form of a straight line up to terms of order $\mathcal{O}(\lambda^3)$. Equation (2.1) is an infinitesimal version of the law of inertia and im-

plies the universality of free fall. It encapsulates some form of the weak equivalence principle as refined by Einstein due to his epiphany by October 1907: "I was sitting in a chair in the patent office at Bern ...". It turns out that λ is an affine parameter of a geodesic. In arbitrary coordinates x^μ, Equation (2.1) takes the form

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \omega^\mu_{\nu\kappa} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\kappa}{\mathrm{d}\lambda} = 0 \tag{2.2}$$

at p, where

$$\omega^\mu_{\nu\kappa} = \frac{\partial x^\mu}{\partial \chi^\rho} \frac{\partial^2 \chi^\rho}{\partial x^\nu \partial x^\kappa}. \tag{2.3}$$

The quantity $\omega^\mu_{\nu\kappa}$ (symmetric in ν, κ) is easily seen that forms the components of a connection and is given by the functions $\chi^\mu\left(x^\nu\right)$. Formally, the analytical expression (2.3) is defined not only at p but at other points as well. However, the correct is that the expression (2.3) is meaningful and gives the values of the connection only at the point p. At another point a corresponding expression will be valid for the corresponding χ^μ, and depending on what assumptions one adopts, the form of the connection will arise at all points. Indeed, if Equation (2.3) was valid in a neighborhood of p, applying this equation for the specific system $x^\mu = \chi^\mu$, it would give a vanishing (flat) connection with vanishing curvature, which certainly is not the situation that one wants to describe. For the same reason, Equation (2.1) cannot be true in a whole neighborhood around p (something that sometimes is misunderstood). However, Equation (2.2) for $x^\mu = \chi^\mu$ provides with the use of (2.1) that at the point p the components of the connection in the system χ^μ vanish. So, basically the system χ^μ defines the connection at p, such that at p the components of the connection in its own coordinates χ^μ vanish.

It is a crucial point that for massive particles the parameter λ of (2.1) has for consistency to be the proper time τ along the orbit of the particle (or more generally equal to $\alpha\tau + \beta$, with α, β constants), defined by $\mathrm{d}\tau^2 = -g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$. This is an assumption inside Equation (2.1), and therefore, for massive particles the only affine parameter is assumed to be the proper time. If this was not assumed, we will see that inconsistency would appear even in our everyday world and the Christoffel connection would never arise. For massless particles there is no proper time, so a parameter λ is assumed to exist such that Equation (2.1) is still valid. In the massless case, the parameter λ can be the temporal coordinate χ^0 since if λ is different than χ^0, one can define the new coordinate system ξ^μ such that its spatial coordinates are $\xi^i = \chi^i$, while $\xi^0 = \lambda$, and then $\frac{\mathrm{d}^2 \xi^\mu}{\mathrm{d}\left(\xi^0\right)^2}(p) = 0$. We will let a general λ satisfying Equation (2.1) without some other specification, except if stated otherwise explicitly. Another point to be noticed is the following: the arbitrary system x^μ can equally well describe the motion of either massive or massless particles (for example, the standard coordi-

nates t,r,θ,ϕ of a spherically symmetric configuration describes all particles). Intuitively, this is so since the orbit of a massless particle is very close to the orbit of a massive particle chasing the first one very closely. Since the connection is a spacetime function, independent of the orbits, the inertial system χ^μ of (2.3) describes all particles.

If, instead of λ, an arbitrary parameter σ is used, then Equation (2.1) becomes

$$\frac{\mathrm{d}^2\chi^\mu}{\mathrm{d}\sigma^2}(p)=f(p)\frac{\mathrm{d}\chi^\mu}{\mathrm{d}\sigma}(p), \tag{2.4}$$

where

$$f=-\frac{\mathrm{d}^2\sigma}{\mathrm{d}\lambda^2}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\lambda}\right)^{-2}, \tag{2.5}$$

while Equation (2.2) becomes

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\sigma^2}+\omega^\mu_{\nu\kappa}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\kappa}{\mathrm{d}\sigma}=f\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma} \tag{2.6}$$

at p. The connection $\omega^\mu_{\nu\kappa}$ remains unchanged when we pass from the affine parameter λ to the non-affine σ. Equation (2.2), Equation (2.6), as known, transform tensorially under coordinate transformations with f unchanged. As for the function f, it could in principle be different from geodesic to geodesic, depending on the selected parametrization on each one.

If a coordinate system $\hat{x}^\mu$ satisfies $\frac{\partial^2\hat{x}^\mu}{\partial\chi^\nu\partial\chi^\kappa}(p)=0$, while the parameter $\hat{\lambda}$ obeys $\frac{\mathrm{d}^2\lambda}{\mathrm{d}\hat{\lambda}^2}(p)=0$, then $\frac{\mathrm{d}^2\hat{x}^\mu}{\mathrm{d}\hat{\lambda}^2}(p)=0$. Thus, it seems that $\chi^\mu(\lambda)$ may not be unique around p with the property that "eats up" gravity, but other systems $\hat{x}^\mu(\hat{\lambda})$ could also exist with the same property. Indeed, for example, the globally rotated systems $\mathcal{X}^\mu=\gamma^\mu_{\ \nu}\chi^\nu$, where $\gamma^\mu_{\ \nu}$ a general invertible constant matrix, are also locally inertial with the same affine parameter λ, and satisfy (2.1), (2.3). In the following of this section, we will show the existence of other non-trivial such systems and construct them more systematically using two methods. In the first method, from an arbitrary system x^μ, another local inertial system $\tilde{x}^\mu$ will be constructed for any sort of particles being present, such that $\tilde{x}^\mu(\lambda)$ satisfies (2.1) and also $\tilde{x}^\mu$ can replace χ^μ in (2.3). In this sense, another system y^μ will define its own local inertial frame $\tilde{y}^\mu$, different in general from $\tilde{x}^\mu$, such that again $\tilde{y}^\mu(\lambda)$ satisfies (2.1), (2.3), and so on. The system $\tilde{y}^\mu(y^\nu)$ can also be expressed in terms of x^μ as $\tilde{y}^\mu(y^\rho(x^\nu))$. The inertial systems and the way they are constructed in the first method do not lead to any inconsistency even if only massless particles are present. They rather convince us in a more direct and manifest manner about the multiplicity of inertial frames and they lead to the Christoffel connection. In the second method, from an arbitrary but specific system x^μ, in the case where only massless particles are present, we will construct an infinity of local inertial frames $\bar{x}^\mu(x^\nu)$ such

that each one of them, along with its appropriate parameter $\bar{\lambda}$, satisfies (2.1), *i.e.* $\bar{x}^\mu\left(\bar{\lambda}\right)$ satisfies (2.1), while it also satisfies a deformation of (2.3). The systems $\bar{x}^\mu$ are different in general from the systems $\chi^\mu, \mathcal{X}^\mu, \tilde{x}^\mu, \tilde{y}^\mu, \cdots$ because the affine parameter $\bar{\lambda}$ is different than λ. If a massive particle is present, the systems $\bar{x}^\mu$ reduce to $\tilde{x}^\mu$, since a $\bar{x}^\mu$ has to describe both massive and massless particles. The existence of the systems $\bar{x}^\mu$ in the purely massless case is the root for the proof of the claimed inconsistency.

For an arbitrary coordinate system x^μ, we define the coordinate system $\tilde{x}^\mu$ by the relation

$$x^\mu - x^\mu(p) = \tilde{x}^\mu - \tilde{x}^\mu(p) - \frac{1}{2}\omega^\mu_{\nu\kappa}(p)\left(\tilde{x}^\nu - \tilde{x}^\nu(p)\right)\left(\tilde{x}^\kappa - \tilde{x}^\kappa(p)\right). \tag{2.7}$$

Maybe a few $\tilde{x}^\mu$ are defined, and not just one, due to the quadratic form of the defining equation. As χ^μ, the system $\tilde{x}^\mu$, due to Equation (2.7), is also valid for all particles. The components $\tilde{\omega}^\mu_{\nu\kappa}$ of the connection in the coordinates $\tilde{x}^\mu$ are easily seen to have $\tilde{\omega}^\mu_{\nu\kappa}(p) = 0$. Therefore, since Equation (2.2), Equation (2.6) hold for any system, applying them for $x^\mu = \tilde{x}^\mu$ we get the corresponding of (2.1), (2.4), *i.e.*

$$\frac{\mathrm{d}^2\tilde{x}^\mu}{\mathrm{d}\lambda^2}(p) = 0 \tag{2.8}$$

$$\frac{\mathrm{d}^2\tilde{x}^\mu}{\mathrm{d}\sigma^2}(p) = f(p)\frac{\mathrm{d}\tilde{x}^\mu}{\mathrm{d}\sigma}(p). \tag{2.9}$$

Moreover, a direct calculation shows that the expression (2.3) is also valid with $\tilde{x}^\mu$ playing the role of χ^μ, *i.e.* it is

$$\omega^\mu_{\nu\kappa} = \frac{\partial x^\mu}{\partial\tilde{x}^\rho}\frac{\partial^2\tilde{x}^\rho}{\partial x^\nu\partial x^\kappa} \tag{2.10}$$

at p.[1] Obviously, due to (2.10), the connection at p defined by a system $\tilde{x}^\mu$ coincides with $\omega^\mu_{\nu\kappa}(p)$. These mean that for each coordinate system x^μ, a corresponding local inertial system $\tilde{x}^\mu$ can be defined with the same affine parameter λ. It can be noticed that Equation (2.7) defining $\tilde{x}^\mu$ can be written equivalently in exactly the same form if x^μ is replaced by $X^\mu = \gamma^\mu_\nu x^\nu$, $\omega^\mu_{\nu\kappa}$ by the corresponding components $\Omega^\mu_{\nu\kappa} = \omega^\rho_{\sigma\tau}\gamma^\mu_\rho\left(\gamma^{-1}\right)^\sigma_\nu\left(\gamma^{-1}\right)^\tau_\kappa$ of the same connection in the X^μ coordinates, and $\tilde{x}^\mu$ by $\tilde{X}^\mu = \gamma^\mu_\nu\tilde{x}^\nu$. The system $\tilde{X}^\mu$ satisfies (2.8), (2.9), while $\Omega^\mu_{\nu\kappa}$, X^μ, $\tilde{X}^\mu$ satisfy (2.10). Therefore, the (trivial) globally rotated frames of $\tilde{x}^\mu$ are locally inertial with the same affine parameter λ and arise from the class of the globally rotated systems of x^μ, still within the formula (2.7). Since Equations (2.8)-(2.10) are analogous to Equation (2.1), Equation (2.4), Equation (2.3), someone might think that the systems $\tilde{x}^\mu$ are not

[1]Inversely, starting from a geodesic equation of the form (2.6) (with f the same for all coordinate systems) for an arbitrary torsionless connection $\omega^\mu{}_{\nu\kappa}$, Equations (2.8)-(2.10) arise at p for the system $\tilde{x}^\mu$ of (2.7). Therefore, if a geodesic equation of some connection provides the orbits of free particles, Equation (2.1) is not anything fundamental, but it is a simple geometric fact that is always true; in the hands of Einstein, however, Equation (2.1) obtained a primitive role which allowed him to determine the connection.

something new, but they coincide with χ^μ (or their rotations $\mathcal{X}^\mu$). However, this is not true in general. Moreover, $\tilde{x}^\mu$ is not the rotation of $\tilde{y}^\mu$ in general. Therefore, various $\tilde{x}^\mu$ exist, which are different from χ^μ and different from each other. To see this, let x^μ defines through (2.7) the system $\tilde{x}^\mu$, while another system y^μ defines $\tilde{y}^\mu$. Let us suppose that $\tilde{x}^\mu$ coincides with $\gamma^\mu{}_\nu \chi^\nu$ and $\tilde{y}^\mu$ coincides with $\Gamma^\mu{}_\nu \chi^\nu$, where $\gamma^\mu{}_\nu, \Gamma^\mu{}_\nu$ are invertible constant matrices. Then, $\tilde{y}^\mu = \Delta^\mu{}_\nu \tilde{x}^\nu$, where $\Delta^\mu{}_\nu = \Gamma^\mu{}_\kappa \left(\gamma^{-1}\right)^\kappa{}_\nu$. This relation between $\tilde{x}^\mu, \tilde{y}^\mu$ is shown not to be generic, even if $\tilde{x}^\mu, \tilde{y}^\mu$ are not related to χ^μ. Indeed, Equation (2.7) is a quadratic algebraic equation for $\tilde{x}^\mu$ and its solution for $\tilde{x}^\mu$ is some algebraic expression containing square roots. Similarly, we solve the second equation with y^μ for $\tilde{x}^\mu$; it will have another form with square roots with some differences due to the different values of the connection in the y^μ system and the possible existence of the free parameters $\Delta^\mu{}_\nu$ wandering around. Equating the two expressions for $\tilde{x}^\mu$ we get that the system y^μ is related to the system x^μ through a specific algebraic expression with square roots, containing also various numerical values. Of course, such an expression does not exhaust an arbitrary coordinate system y^μ, so the above relation of $\tilde{x}^\mu$, $\tilde{y}^\mu$ only rarely occurs. Therefore, formula (2.7) defines the first method mentioned above for the determination of various inertial frames.

There are two non-trivial things happening for the geodesic Equation (2.6) with an arbitrary $\omega^\mu_{\nu\kappa}$. First, Equation (2.6) is quasi-form invariant under changes of the parameter σ, something that does not happen in all ordinary differential equations, although a change of the independent variable is always a permissible transformation. This precisely means that the form of Equation (2.6) does not change if σ is replaced by $\breve{\sigma}$ and f by $\breve{f}$, where $\breve{f} = \left(f \frac{\mathrm{d}\breve{\sigma}}{\mathrm{d}\sigma} - \frac{\mathrm{d}^2\breve{\sigma}}{\mathrm{d}\sigma^2} \right) \left(\frac{\mathrm{d}\breve{\sigma}}{\mathrm{d}\sigma} \right)^{-2}$ (and to be more precise x^μ by $x^\mu \circ \sigma$). Therefore, when we have a connection $\omega^\mu_{\nu\kappa}$ at hand (in a coordinate system), the choice of the function f does not affect the geodesics (their point sets). So, while the new geodesic equation with $\breve{f}$ to be solved is different than the original one with f, the geodesics are the same. The above expression for $\breve{f}$ is in agreement with (2.5) setting $\breve{f} = 0$ and $\breve{\sigma} = \lambda$.

Second, if $\overline{q}^\mu$ is an arbitrary vector field, then

$$\underset{\sim}{\omega}{}^\mu_{\nu\kappa} = \omega^\mu_{\nu\kappa} + \delta^\mu_\nu \overline{q}_\kappa + \delta^\mu_\kappa \overline{q}_\nu \tag{2.11}$$

defines another torsionless connection. The value of $\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$ at p depends on the value of $\overline{q}^\mu$ at p and no differentiation of $\overline{q}^\mu$ or $\omega^\mu_{\nu\kappa}$ will appear. However, the values of $\overline{q}^\mu$ along the geodesic will be used in order to make changes in the parameter σ or λ, similarly to the fact that the various coordinate systems are also defined around p. The geodesic Equation (2.6) for $\omega^\mu_{\nu\kappa}$ is written as

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\sigma^2} + \underset{\sim}{\omega}{}^\mu_{\nu\kappa} \frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma} \frac{\mathrm{d}x^\kappa}{\mathrm{d}\sigma} = \overline{f} \frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma} \tag{2.12}$$

at p, where

$$\bar{f} = f + 2\bar{q}_\rho \frac{\mathrm{d}x^\rho}{\mathrm{d}\sigma}. \tag{2.13}$$

Moreover, defining $\bar{\sigma}$ by the equation

$$\frac{\mathrm{d}^2\bar{\sigma}}{\mathrm{d}\sigma^2} + f\left(\frac{\mathrm{d}\bar{\sigma}}{\mathrm{d}\sigma}\right)^2 - \bar{f}\frac{\mathrm{d}\bar{\sigma}}{\mathrm{d}\sigma} = 0, \tag{2.14}$$

Equation (2.12) becomes

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\bar{\sigma}^2} + \underset{\sim}{\omega}{}^\mu_{\nu\kappa}\frac{\mathrm{d}x^\nu}{\mathrm{d}\bar{\sigma}}\frac{\mathrm{d}x^\kappa}{\mathrm{d}\bar{\sigma}} = f\frac{\mathrm{d}x^\mu}{\mathrm{d}\bar{\sigma}} \tag{2.15}$$

at p. Alternatively, Equation (2.15) can arise by first performing a parameter change from σ to $\bar{\sigma}$ in (2.6) and then introducing $\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$ from (2.11). In the context of geometry, the previous equations for $\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$ are valid not only at p as here, but along the whole curves, and then, since (2.12) is also the geodesic equation of $\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$, this means that the geodesics of $\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$, $\omega^\mu_{\nu\kappa}$ coincide ($\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$ is called projectively equivalent to $\omega^\mu_{\nu\kappa}$). So, two projectively equivalent connections have different f's with the same parametrization (*i.e.* σ), or alternatively, they can have the same f with different parametrizations (*i.e.* $\sigma, \bar{\sigma}$). In practice, Equation (2.15) for suitable $\bar{q}^\mu, \bar{\sigma}$ can facilitate the finding of the geodesics. Hence, the connection $\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$ defined by (2.11) is in some loose sense a sort of "gauge" transformation for the geodesics. For our purposes here, the validity of the previous equations at p will be enough. The inverse is also true, which means that any two connections which have all their geodesics the same are necessarily related by an equation of the form (2.11) [28] [29]. Indeed, if $v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}$, the geodesic equations of the projectively equivalent connections $\omega^\mu_{\nu\kappa}, \underset{\sim}{\omega}{}^\mu_{\nu\kappa}$ are written as $\frac{\mathrm{d}v^\mu}{\mathrm{d}\sigma} + \omega^\mu_{\nu\kappa}v^\nu v^\kappa = hv^\mu$, $\frac{\mathrm{d}v^\mu}{\mathrm{d}\sigma} + \underset{\sim}{\omega}{}^\mu_{\nu\kappa}v^\nu v^\kappa = \underset{\sim}{h}v^\mu$ for some appropriate functions $h, \underset{\sim}{h}$. Therefore $\bar{\Delta}_{\mu\nu\kappa}v^\nu v^\kappa = Hv_\mu$, where $\bar{\Delta}_{\mu\nu\kappa} = \omega_{\mu\nu\kappa} - \underset{\sim}{\omega}{}_{\mu\nu\kappa} = \bar{\Delta}_{\mu\kappa\nu}$, $H = h - \underset{\sim}{h}$. The above algebraic condition for v^μ implies that there are vector fields A^μ, B^μ, such that $\bar{\Delta}_{\mu\nu\kappa} = g_{\mu\nu}A_\kappa + g_{\mu\kappa}B_\nu$. The symmetry property of $\bar{\Delta}_{\mu\nu\kappa}$ implies $A_\mu = B_\mu$, and thus Equation (2.11) arises.

We now focus to the case where only massless particles are present. We will find other local inertial frames, for which the equivalence principle holds, in a different way. Namely, using a single coordinate system x^μ, we will define various systems $\bar{x}^\mu$ (other than the global rotations of the corresponding system $\tilde{x}^\mu$) such that still gravity is locally eliminated. If in particular we attach with such a $\bar{x}^\mu$ the appropriate affine parameter $\bar{\lambda}$, we will have the form (2.1) for $\bar{x}^\mu\left(\bar{\lambda}\right)$. Indeed, we have defined in (2.14) the parameter $\bar{\sigma}$, such that the geodesic equation for $\underset{\sim}{\omega}{}^\mu_{\nu\kappa}$ at p takes the specific form (2.15). Define now the parameter $\bar{\lambda}$ by the equation

$$f = \frac{\mathrm{d}^2\bar{\lambda}}{\mathrm{d}\bar{\sigma}^2}\left(\frac{\mathrm{d}\bar{\lambda}}{\mathrm{d}\bar{\sigma}}\right)^{-1} \tag{2.16}$$

and we get

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\bar{\lambda}^2} + \underset{\sim}{\omega}^\mu{}_{\nu\kappa} \frac{\mathrm{d}x^\nu}{\mathrm{d}\bar{\lambda}} \frac{\mathrm{d}x^\kappa}{\mathrm{d}\bar{\lambda}} = 0 \tag{2.17}$$

at p. The coordinate system $\bar{x}^\mu$ defined by

$$x^\mu - x^\mu(p) = \bar{x}^\mu - \bar{x}^\mu(p) - \frac{1}{2}\underset{\sim}{\omega}^\mu{}_{\nu\kappa}(p)\left(\bar{x}^\nu - \bar{x}^\nu(p)\right)\left(\bar{x}^\kappa - \bar{x}^\kappa(p)\right) \tag{2.18}$$

has vanishing components at p of the $\underset{\sim}{\omega}$ connection in the $\bar{x}^\mu$ system, *i.e.* $\underset{\sim}{\bar{\omega}}^\mu{}_{\nu\kappa}(p) = 0$, quite similarly to what happens with Equation (2.7). Due to (2.18), the system $\bar{x}^\mu$ is valid for all particles, massive and massless, similarly to what happens with the systems $\tilde{x}^\mu$. Since Equation (2.17) holds for any system, applying it for $x^\mu = \bar{x}^\mu$, we get the corresponding of (2.1), *i.e.*

$$\frac{\mathrm{d}^2 \bar{x}^\mu}{\mathrm{d}\bar{\lambda}^2}(p) = 0. \tag{2.19}$$

Moreover, it arises

$$\underset{\sim}{\omega}^\mu{}_{\nu\kappa} = \frac{\partial x^\mu}{\partial \bar{x}^\rho} \frac{\partial^2 \bar{x}^\rho}{\partial x^\nu \partial x^\kappa} \tag{2.20}$$

at p. These mean that various local inertial coordinate systems $\bar{x}^\mu$, each one with an appropriate affine parameter $\bar{\lambda}$, arise from the projectively equivalent connections $\underset{\sim}{\omega}^\mu{}_{\nu\kappa}$ (as parametrized by $\bar{q}^\mu$). Due to (2.20), the connection at p defined by a system $\bar{x}^\mu$ coincides with $\underset{\sim}{\omega}^\mu{}_{\nu\kappa}(p)$.

Since the role played by the parameter $\bar{\lambda}$ is crucial, it is enlightening to simplify the previous analysis in order to make the situation even more clear. We consider in (2.12) the affine parameter λ instead of the general parameter σ and get

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \underset{\sim}{\omega}^\mu{}_{\nu\kappa} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\kappa}{\mathrm{d}\lambda} = 2\bar{q}_\rho \frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \tag{2.21}$$

at p. Equation (2.21) also arises directly from (2.2) introducing the quantity $\underset{\sim}{\omega}^\mu{}_{\nu\kappa}$. Then, we define directly the parameter $\bar{\lambda}$ from the equation

$$\frac{\mathrm{d}^2 \bar{\lambda}}{\mathrm{d}\lambda^2} = 2\bar{q}_\rho \frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda} \frac{\mathrm{d}\bar{\lambda}}{\mathrm{d}\lambda}. \tag{2.22}$$

Equation (2.22) also arises from the previous equations, since for $f = 0$ in (2.16), a solution can be $\bar{\lambda} = \bar{\sigma}$, and Equation (2.14) provides (2.22). Finally, changing in (2.21) from λ to $\bar{\lambda}$ according to (2.22), we get (2.17). We can summarize by saying that adding in (2.2) the $\bar{q}_\mu$ terms in order to get the projectively equivalent connection $\underset{\sim}{\omega}^\mu{}_{\nu\kappa}$ as in (2.21), a non-affine geodesic arises which can be changed into the affine one (2.17) through (2.22). Equations (2.19), (2.20) follow exactly as before. The parameter $\bar{\lambda}$ is different from λ (2.or $\alpha\lambda + \beta$ with α, β constants), otherwise Equation (2.22) would be inconsistent. So, the process is that a $\bar{q}^\mu$ defines a $\underset{\sim}{\omega}^\mu{}_{\nu\kappa}$ through (2.11) and a $\bar{\lambda}$ through (2.22), then a $\bar{x}^\mu$ is defined through (2.18) and satisfies Equation (2.19), Equation (2.20). The above description shows that instead of changing the coordinate

systems to get other local inertial frames, we can stay at a single coordinate system and just flip to projectively equivalent connections in order to get new local inertial frames. Therefore, formula (2.18) defines the second method mentioned above for the determination of inertial frames in the purely massless case.

As in the case with the systems $\tilde{x}^\mu$, also here, Equation (2.18) defining $\bar{x}^\mu$ can be written equivalently in exactly the same form if x^μ is replaced by $X^\mu = \gamma^\mu_{\ \nu} x^\nu$, $\underset{\sim}{\omega}^\mu_{\ \nu\kappa}$ by the corresponding components $\underset{\sim}{\Omega}^\mu_{\ \nu\kappa} = \underset{\sim}{\omega}^\rho_{\ \sigma\tau}\gamma^\mu_{\ \rho}\left(\gamma^{-1}\right)^\sigma_{\ \nu}\left(\gamma^{-1}\right)^\tau_{\ \kappa}$ of the $\underset{\sim}{\omega}$ connection in the X^μ coordinates, and $\bar{x}^\mu$ by $\bar{X}^\mu = \gamma^\mu_{\ \nu}\bar{x}^\nu$. The system $\bar{X}^\mu$ satisfies (2.19), while $\underset{\sim}{\Omega}^\mu_{\ \nu\kappa}$, X^μ, $\bar{X}^\mu$ satisfy (2.20). Therefore, the (trivial) globally rotated frames of $\bar{x}^\mu$ are locally inertial with affine parameter $\bar{\lambda}$ and arise from the class of the globally rotated systems of x^μ, still within the formula (2.18). The systems $\bar{x}^\mu$ are different in general (i) from the systems $\tilde{x}^\mu$ or $\tilde{X}^\mu$, and (ii) from $\mathcal{X}^\mu$ and from each other. As for (i), this is seen by combining Equation (2.7), Equation (2.18) and $\tilde{x}^\mu = \gamma^\mu_{\ \nu}\bar{x}^\nu$, from where it arises $\left[\gamma^\mu_{\ \rho}\omega^\rho_{\ \nu\kappa}(p) - \underset{\sim}{\Omega}^\mu_{\ \nu\kappa}(p)\right]\left[\tilde{x}^\nu - \tilde{x}^\nu(p)\right]\left[\tilde{x}^\kappa - \tilde{x}^\kappa(p)\right] = 2\left(\gamma^\mu_{\ \nu} - \delta^\mu_{\ \nu}\right)\left[\tilde{x}^\nu - \tilde{x}^\nu(p)\right]$. There is no $\gamma^\mu_{\ \nu}$ such that the last equation is valid, since the left hand side is quadratic in $\tilde{x}^\mu - \tilde{x}^\mu(p)$ while the right hand side is linear, and also $\tilde{x}^\mu$ depends on x^μ. The same fact can also be seen as follows. The solution of Equation (2.22) will in general have both $\frac{\mathrm{d}^2\bar{\lambda}}{\mathrm{d}\lambda^2}(p)$, $\frac{\mathrm{d}\bar{\lambda}}{\mathrm{d}\lambda}(p)$ non-vanishing, since the opposite is a very special initial condition for the differential Equation (2.22), which means a very special choice for the solution $\bar{\lambda}(\lambda)$. If $\tilde{x}^\mu = \gamma^\mu_{\ \nu}\bar{x}^\nu$, then Equation (2.19), (2.8) would give $\frac{\mathrm{d}^2\bar{\lambda}}{\mathrm{d}\lambda^2}(p) = 0$, which is incompatible in general. As for (ii), let a $\bar{q}^\mu$ defines $\underset{\sim}{\omega}^\mu_{\ \nu\kappa}$ and $\bar{x}^\mu$, while another $\bar{\bar{q}}^\mu$ defines $\underset{\sim}{\omega}^\mu_{\ \nu\kappa}$ and $\bar{\bar{x}}^\mu$. If $\bar{x}^\mu = \gamma^\mu_{\ \nu}\chi^\nu$ and $\bar{\bar{x}}^\mu = \Gamma^\mu_{\ \nu}\chi^\nu$, then $\bar{\bar{x}}^\mu = \Delta^\mu_{\ \nu}\bar{x}^\nu$, where $\Delta^\mu_{\ \nu} = \Gamma^\mu_{\ \kappa}\left(\gamma^{-1}\right)^\kappa_{\ \nu}$. This relation between $\bar{x}^\mu, \bar{\bar{x}}^\mu$ is not generic, even if $\bar{x}^\mu, \bar{\bar{x}}^\mu$ are not related to χ^μ. Indeed, from Equation (2.18) applied twice for $\bar{x}^\mu, \bar{\bar{x}}^\mu$, we get $\left[\underset{\sim}{\omega}^\mu_{\ \rho\sigma}(p)\Delta^\rho_{\ \nu}\Delta^\sigma_{\ \kappa} - \underset{\sim}{\omega}^\mu_{\ \nu\kappa}(p)\right]\left[\bar{x}^\nu - \bar{x}^\nu(p)\right]\left[\bar{x}^\kappa - \bar{x}^\kappa(p)\right]$ $= 2\left(\Delta^\mu_{\ \nu} - \delta^\mu_{\ \nu}\right)\left[\bar{x}^\nu - \bar{x}^\nu(p)\right]$. As before, there is no $\Delta^\mu_{\ \nu}$ such that the last equation is valid, since the left hand side is quadratic in $\bar{x}^\mu - \bar{x}^\mu(p)$ while the right hand side is linear, and also $\bar{x}^\mu$ depends on x^μ. The same conclusion is also derived from Equation (2.22) applied twice for $\bar{\lambda}, \bar{\bar{\lambda}}$, and we get $\frac{\mathrm{d}^2\bar{\bar{\lambda}}}{\mathrm{d}\bar{\lambda}^2}\left(\frac{\mathrm{d}\bar{\lambda}}{\mathrm{d}\lambda}\right)^2 = 2\left(\bar{\bar{q}}_\rho - \bar{q}_\rho\right)\frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda}\frac{\mathrm{d}\bar{\bar{\lambda}}}{\mathrm{d}\lambda}$. If $\bar{\bar{x}}^\mu = \Delta^\mu_{\ \nu}\bar{x}^\nu$, then Equation (2.19) applied twice for $\bar{x}^\mu, \bar{\bar{x}}^\mu$ provides $\frac{\mathrm{d}^2\bar{\bar{\lambda}}}{\mathrm{d}\bar{\lambda}^2}(p) = 0$, and therefore $\bar{\bar{q}}^\mu(p) = \bar{q}^\mu(p)$, which is inconsistent in general.

When massive particles are not present, the root of the claimed inconsistency lies on the existence of the local inertial frames $\bar{x}^\mu$ which give rise to the pro-

jectively equivalent connections $\underset{\sim}{\omega}^{\mu}{}_{\nu\kappa}$ according to (2.20). If, however, massive particles are present, the existence of proper time for the massive particles changes completely the situation and the systems $\bar{x}^{\mu}$ reduce to $\tilde{x}^{\mu}$. Equation (2.17), Equation (2.19) for the massive particles should have $\bar{\lambda} = \lambda = \tau$ (or $\alpha\tau + \beta$) since the weak equivalence principle in the form (2.1) has embedded the assumption of proper time. Although the affine parameters between massive and massless particles are different, the system $\bar{x}^{\mu}$ should be the same for both. Equation (2.22) for a massive particles provides $\bar{q}^{\mu} = 0$, thus $\underset{\sim}{\omega}^{\mu}{}_{\nu\kappa} = \omega^{\mu}{}_{\nu\kappa}$ and $\bar{x}^{\mu} = \tilde{x}^{\mu}$. Therefore, if massive particles are present, the Christoffel connection arises normally and no inconsistency occurs. The same conclusion arises, although in a little more complicated way, if we think the presence of massive particles in terms of the general non-affine parameter σ instead of the proper time. Indeed, Equation (2.16) defines the parameter $\bar{\sigma}(\tau)$. Since $\bar{\lambda} = \lambda = \tau$, Equation (2.16), Equation (2.5) coincide, and thus $\bar{\sigma} = \sigma$ (up to a possible additive constant). Hence, from Equation (2.14) it arises $\bar{f} = f$ and (2.13) gives $\bar{q}^{\mu} = 0$. In other words, a general $\underset{\sim}{\omega}^{\mu}{}_{\nu\kappa}$, different than $\omega^{\mu}{}_{\nu\kappa}$, satisfies for massive particles Equation (2.21) with a non-vanishing right hand side, in contrast to Equation (2.2), which means that $\bar{x}^{\mu}$ satisfies (2.19) and not the axiom of the weak equivalence principle for massive particles in the form (2.1). Finally, let us finish with a hypothetical comment: If $\bar{x}^{\mu}$ for timelike particles was different than $\bar{x}^{\mu}$ for null particles (which is not the case), then the construction of the systems $\bar{x}^{\mu}$ would invalidate the derivation of the Christoffel connection even in the presence of massive particles.

3. The Local Lorentz Frames

Given a spacetime metric $g_{\mu\nu}$, for a given point p there is a surrounding coordinate system x'^{μ}, such that the corresponding components have $g'_{\mu\nu}(p) = \eta_{\mu\nu}$. This is an issue of linear algebra since the matrix $g_{\mu\nu}(p)$ can always be diagonalized and x'^{μ} can be constructed such that the vectors $\partial/\partial x'^{\mu}$ at p coincide with the orthonormal vectors of diagonalization. Therefore, an infinite number of such systems x'^{μ} exist. Indeed, any system x''^{μ}, such that $x''^{\mu} - x''^{\mu}(p)$ equals $x'^{\mu} - x'^{\mu}(p)$ plus quadratic or higher powers of $x'^{\mu} - x'^{\mu}(p)$, has $g''_{\mu\nu}(p) = \eta_{\mu\nu}$. While the weak equivalence principle intents to embody gravity, a metric with spacetime signature intends to embrace special relativity. Although the local Minkowski structure is always present (basically through the tangent space) and a coordinate system exists such that the metric locally coincides with the Minkowski metric, obviously there are coordinate systems around p with $g_{\mu\nu}(p) \neq \eta_{\mu\nu}$. In an arbitrary coordinate system x^{μ} it holds

$$g_{\mu\nu} = \frac{\partial x'^{\rho}}{\partial x^{\mu}} \frac{\partial x'^{\sigma}}{\partial x^{\nu}} \eta_{\rho\sigma} \tag{3.1}$$

at p. This equation cannot be true in a whole neighborhood around p, since then, the space would be Minkowski in a whole neighborhood, which certainly is not the situation that one wants to describe. Today, a metric $g_{\mu\nu}$ is considered as a

fundamental object, but it was not always obvious that such a $g_{\mu\nu}$ should exist and generalize the Minkowski metric in the presence of gravity. It seems that what is called Einstein equivalence principle implies the existence of a $g_{\mu\nu}$ which is locally Minkowski[2]. This principle extends the weak equivalence principle and requires the independency from the direction and spacetime position for any freely falling frame (where inhomogeneities of the external fields are ignorable) of any local non-gravitational test experiment. Thus, the metric arises as a direct result of the existence of the freely falling frames together with their direction insensitivity (local Lorentz invariance of the non-gravitational laws of physics) and position independency (see [32], p. 23 for more details). However, the selection of the Christoffel connection is not automatic and has to be justified.

Another coordinate system $\mathring{x}'^{\mu}$, more specific than x'^{μ}, can be defined through the relation

$$x'^{\mu}-x'^{\mu}(p)=\mathring{x}'^{\mu}-\mathring{x}'^{\mu}(p)-\frac{1}{2}\Gamma'^{\mu}{}_{\nu\kappa}(p)\left(\mathring{x}'^{\nu}-\mathring{x}'^{\nu}(p)\right)\left(\mathring{x}'^{\kappa}-\mathring{x}'^{\kappa}(p)\right), \quad (3.2)$$

where $\Gamma'^{\mu}{}_{\nu\kappa}$ are the components of the Christoffel connection in the system x'^{μ}. Then, it is easily seen that $\mathring{g}'_{\mu\nu}(p)=\eta_{\mu\nu}$ and $\mathring{g}'_{\mu\nu}\left(\mathring{x}'^{\kappa}\right)$ has vanishing first derivatives at *p*, *i.e.* $\frac{\partial\mathring{g}'_{\mu\nu}}{\partial\mathring{x}'^{\kappa}}(p)=0$, $\mathring{\Gamma}'^{\mu}{}_{\nu\kappa}(p)=0$ (the small circle above the corresponding symbols reminds us notationally of these zero values). Also, Equation (3.1) holds at *p* with $\mathring{x}'^{\mu}$ taking the place of x'^{μ}. These are simple geometrical facts which are generally true and are sometimes useful for performing proofs in tensor analysis. A system $\mathring{x}'^{\mu}$ is called local Lorentz frame (or local Minkowski frame) around *p*. Therefore, gravitation, which is manifested through $g_{\mu\nu}$ (and possibly other fields), is a second-order effect. Certainly the existence of coordinate systems such as x'^{μ}, or even more $\mathring{x}'^{\mu}$, expresses the fact that a gravity theory agrees locally with special relativity to a good approximation. In an arbitrary coordinate system x^{μ} it is

$$g_{\mu\nu}=\frac{\partial\mathring{x}'^{\rho}}{\partial x^{\mu}}\frac{\partial\mathring{x}'^{\sigma}}{\partial x^{\nu}}\mathring{g}'_{\rho\sigma}. \quad (3.3)$$

Differentiation of (3.3) gives

$$\frac{\partial g_{\mu\nu}}{\partial x^{\kappa}}=\eta_{\rho\sigma}\frac{\partial\mathring{x}'^{\rho}}{\partial x^{\mu}}\frac{\partial^2\mathring{x}'^{\sigma}}{\partial x^{\nu}\partial x^{\kappa}}+\eta_{\rho\sigma}\frac{\partial\mathring{x}'^{\rho}}{\partial x^{\nu}}\frac{\partial^2\mathring{x}'^{\sigma}}{\partial x^{\mu}\partial x^{\kappa}} \quad (3.4)$$

at *p*.

If we consider the (global) Lorentz transformation $x'^{\mu}=\Lambda^{\mu}_{\nu}x'^{\nu}$ of x'^{μ}, then the corresponding metric $g'_{\mu\nu}$ also has the property $g'_{\mu\nu}(p)=\eta_{\mu\nu}$. Accordingly, the systems $\mathring{x}'^{\mu}=\Lambda^{\mu}_{\nu}\mathring{x}'^{\nu}$ satisfy $\mathring{g}'_{\mu\nu}(p)=\eta_{\mu\nu}$, $\frac{\partial\mathring{g}'_{\mu\nu}}{\partial\mathring{x}'^{\kappa}}(p)=0$, so they define other local Lorentz frames. Equation (3.2) defining $\mathring{x}'^{\mu}$ can be written equiva-

[2]The Hughes-Drever experiment rules out the existence of more than one second-rank tensor field both coupling directly to matter, however, vector and tensor fields which couple only to gravity or to matter's self gravitational energy are not ruled out [30] [31].

lently in exactly the same form if x'^{μ} is replaced by $\varkappa'^{\mu}$, $\Gamma'^{\mu}_{\nu\kappa}$ by the corresponding components $\Gamma'^{\rho}_{\sigma\tau}\Lambda^{\mu}_{\rho}\left(\Lambda^{-1}\right)^{\sigma}_{\nu}\left(\Lambda^{-1}\right)^{\tau}_{\kappa}$ of the Christoffel connection in the $\varkappa'^{\mu}$ coordinates, and $\overset{\circ}{x}{}'^{\mu}$ by $\overset{\circ}{\varkappa}{}'^{\mu}$. Therefore, the globally rotated systems of $\overset{\circ}{x}{}'^{\mu}$ arise from the class of the globally rotated systems of x'^{μ}, still within the formula (3.2). Equation (3.3), Equation (3.4) remain the same with $\overset{\circ}{\varkappa}{}'^{\mu}$ substituting $\overset{\circ}{x}{}'^{\mu}$.

4. The Proof of the Inconsistency

The notion of connection (which determines the free fall of particles) and the notion of metric (which embodies the local Minkowski structure) are in general unrelated. It was Einstein's ingenuity to relate them in order to determine the connection $\omega^{\mu}_{\nu\kappa}$. So, the Christoffel connection arises by the identification of a freely falling frame χ^{μ} or $\tilde{x}^{\mu}$ with a local Lorentz frame $\overset{\circ}{x}{}'^{\mu}$, *i.e.* $\chi^{\mu}=\overset{\circ}{x}{}'^{\mu}$ or $\tilde{x}^{\mu}=\overset{\circ}{x}{}'^{\mu}$. Substituting the second derivatives in (3.4) from (2.3) or (2.10), and making use of (3.1) for $\overset{\circ}{x}{}'^{\mu}$, it arises

$$\frac{\partial g_{\mu\nu}}{\partial x^{\kappa}}=g_{\mu\sigma}\omega^{\sigma}_{\nu\kappa}+g_{\nu\sigma}\omega^{\sigma}_{\mu\kappa} \tag{4.1}$$

at *p*. This is the known equation of vanishing non-metricity and for a torsionless connection it gives, after some algebraic manipulation, the Christoffel connection,

$$\omega^{\mu}_{\nu\kappa}=\Gamma^{\mu}_{\nu\kappa} \tag{4.2}$$

at *p*. The connection was found to be the Christoffel connection at the point *p*, and since the same can be repeated for any point, finally the Christoffel connection arises globally.

If massive particles are present (possibly together with massless), no systems $\bar{x}^{\mu}$ exist other than $\tilde{x}^{\mu}$, as explained in Section 2, due to the existence of proper time assumed in the weak equivalence principle for massive particles. Thus, nothing more can be said beyond Equation (4.2) which is the final word for the connection. For a massive particle with $\lambda=\tau$ we have the obvious equation $\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(g_{\mu\nu}\frac{\mathrm{d}x^{\mu}}{\mathrm{d}\lambda}\frac{\mathrm{d}x^{\nu}}{\mathrm{d}\lambda}\right)=0$. The left hand side of this equation, after making use of (2.2) for the connection (4.2), turns out to be identically zero. This is a consistency check for the assumption $\lambda=\tau$ in (2.1).

On the other hand, if only massless particles are present, the freely falling frames χ^{μ}, $\tilde{x}^{\mu}$ still exist, but now the extra local inertial frames $\bar{x}^{\mu}$ are also present. Each one of all these systems should be identified with some local Lorentz frame $\overset{\circ}{x}{}'^{\mu}$. For $\chi^{\mu}=\overset{\circ}{x}{}'^{\mu}$ or $\tilde{x}^{\mu}=\overset{\circ}{x}{}'^{\mu}$, Equation (4.1), Equation (4.2) at *p* arise in exactly the same way as before. However, the unavoidable identification of $\bar{x}^{\mu}$ with some $\overset{\circ}{x}{}'^{\mu}$, *i.e.* $\bar{x}^{\mu}=\overset{\circ}{x}{}'^{\mu}$, gives some extra conditions. Indeed, substituting the second derivatives in (3.4) from (2.20), and making use of (3.1) for $\overset{\circ}{x}{}'^{\mu}$, it arises

$$\frac{\partial g_{\mu\nu}}{\partial x^{\kappa}}=g_{\mu\sigma}\underset{\sim}{\omega}{}^{\sigma}_{\nu\kappa}+g_{\nu\sigma}\underset{\sim}{\omega}{}^{\sigma}_{\mu\kappa} \tag{4.3}$$

at p. Again, algebraic manipulation of Equation (4.3), as before, provides that $\underset{\sim}{\omega}^{\mu}{}_{\nu\kappa}$ is the Christoffel connection, *i.e.*

$$\underset{\sim}{\omega}^{\mu}{}_{\nu\kappa} = \Gamma^{\mu}{}_{\nu\kappa} \tag{4.4}$$

at p. Combining Equation (4.2), Equation (4.4) gives $\delta^{\mu}_{\nu}\overline{q}_{\kappa} + \delta^{\mu}_{\kappa}\overline{q}_{\nu} = 0$ at p. Multiplication of this equation with $\overline{q}_{\mu}$ gives $\overline{q}_{\nu}\overline{q}_{\kappa} = 0$ at p. This last equation obviously gives $\overline{q}^{\mu} = 0$ at p. To get this result even more manifestly, since $\overline{q}^{\mu}$ is arbitrary, there are various vector fields $\overline{q}^{\mu}$ with non-vanishing magnitude at p, *i.e.* $\overline{q}^{\mu}\overline{q}_{\mu} \neq 0$ at p. Considering equation $\overline{q}_{\nu}\overline{q}_{\kappa} = 0$ for such a $\overline{q}^{\mu}$ and multiplying with $\overline{q}^{\kappa}$, we get $\overline{q}_{\nu} = 0$. Equation $\overline{q}_{\nu} = 0$ at p is certainly inconsistent since $\overline{q}^{\mu}$ is arbitrary. Therefore, we have proved that the identification of a local inertial frame with a local Lorentz frame in the purely massless case is meaningless and does not provide any connection (Christoffel or other). The reason for this inconsistency is the absence of proper time and the existence of a distinct affine parameter $\overline{\lambda}$, a different one for each local inertial frame $\overline{x}^{\mu}$, which satisfies Equation (2.19). The local inertial systems $\overline{x}^{\mu}\left(\overline{\lambda}\right)$ are equally good as the local inertial systems $\chi^{\mu}\left(\lambda\right)$ or $\tilde{x}^{\mu}\left(\lambda\right)$ which satisfy Equation (2.1), Equation (2.8). The infinite number of the local inertial systems $\overline{x}^{\mu}$ arise from the infinite number of the projectively equivalent connections $\underset{\sim}{\omega}^{\mu}{}_{\nu\kappa}$, which are parametrized by the vector field $\overline{q}^{\mu}$. The inconsistency is due to that any such connection turns out to coincide with the Christoffel connection. If for massive particles it was not assumed that the affine parameter is the proper time, then it becomes obvious that the same inconsistency would occur along Equation (4.3), Equation (4.4) even in the presence of massive particles.

5. Conclusions

The assumption of the existence of a local inertial coordinate system around a spacetime point, for which the weak equivalence principle is valid, automatically implies the existence of infinitely many other such systems. The identification of any such system with a local Lorentz coordinate system implies that the Christoffel connection controls the kinematics of all particles. However, if only massless particles are present, the situation changes drastically and we show that the identification of a freely falling frame with a local Lorentz frame is meaningless and no connection arises. The presence of the conformal structure for massless particles, instead of the Lorentz one, does not influence our result.

The reason for this inconsistency is that the absence of proper time in the purely massless case allows for the existence of even more local inertial coordinate systems $\overline{x}^{\mu}$. Technically, these last systems arise from the existence of projectively equivalent connections $\underset{\sim}{\omega}^{\mu}{}_{\nu\kappa}$ through the expression (2.18) and are parametrized by a vector field $\overline{q}^{\mu}$ along the formula (2.11). The appropriate affine parameter $\overline{\lambda}$ of these systems for the null particles is defined from (2.22) and Equation (2.19) of free motion is derived, as well as the corresponding geodesic Equation (2.17). The inconsistency arises from Equation (2.20), when it is attempted to identify $\overline{x}^{\mu}$ with a local Lorentz frame, since all the connections

$\tilde{\omega}^{\mu}{}_{\nu\kappa}$ turn out to coincide with the Christoffel one.

Christoffel connection is theoretically significant in physics not because it is introduced as an axiom, or because it is economical and solely constructed in terms of the metric, but because it arises as a result of basic ideas concerning particle kinematics. If there are spacetime regions where only massless particles exist and no massive particles are present to move in the interior of the null cone, the Christoffel connection cannot arise out of the above particle kinematics, and therefore, this connection probably loses its overall significance as an exact connection of spacetime. A connection arising from different ideas of particle kinematics, which are meaningful in the whole of spacetime, is probably more satisfactory conceptually.

In the early universe, before spontaneous symmetry breaking at electroweak scale, all particles are expected to be massless. In this region, which is not expected to be the full quantum gravity regime, notions such as orbits are still meaningful classically or semiclassically. The massless particles in this spacetime region have to decide about their motion in a gravitational environment, but they cannot be based on the notion of a massive particle which appears later in time. Therefore, the kinematical emergence of Christoffel connection in this region breaks down and can set in doubt its overall reliability. Other spacetime regions with only massless particles could be imagined, although probably with no strong arguments, such as the very interior of astrophysical objects or black holes, or relevant objects in the presence of extra dimensions.

Let us finish with an interesting remark. Suppose we have a gravity theory where a connection (or some of its components) is part of the dynamical fields. There are three options. In the first, one insists, due to Einstein's kinematical arguments, that the motion of particles is still governed by the Christoffel connection and that the other connection only indirectly influences the orbits through its interaction with the metric. However, this reasoning is insufficient whenever the previously shown inconsistency of the derivation of Christoffel connection is valid. In the second option, the orbit is governed by the geodesic equation of the connection which caries torsional or non-metricity degrees of freedom. In general, this equation will not preserve in time the nullity constraint of a massless particle, except if this demand is appropriately taken into account, possibly in the construction of the connection. In the third option, the motion of the particles is provided by some action. This action can either refer to extended fields which give the equation of motion for the particle through the geometric optics limit, or it can be a delta-like action providing the orbit directly. Beyond that there are many correction terms that can be added in such actions, similarly to the second case, the respect to the nullity condition only occasionally is expected to occur, if not built-in inside the action.

Acknowledgements

I would like to thank I. Dalianis, A. Kehagias, E. Kiritsis, M. Maggiore, T. Sotiriou, T. Tomaras, R. Troncoso and V. Zarikas for useful discussions.

Conflicts of Interest

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

References

[1] Schmidt, H.-J. (2007) *International Journal of Geometric Methods in Modern Physics*, **4**, 209. https://doi.org/10.1142/S0219887807001977

[2] Sotiriou, T.P. and Faraoni, V. (2010) *Reviews of Modern Physics*, **82**, 451. https://doi.org/10.1103/RevModPhys.82.451

[3] De Felice, A. and Tsujikawa, S. (2010) *Living Reviews in Relativity*, **13**, 3. https://doi.org/10.12942/lrr-2010-3

[4] Nojiri, S.I. and Odintsov, S.D. (2011) *Physics Reports*, **505**, 59. https://doi.org/10.1016/j.physrep.2011.04.001

[5] Madsen, M.S. and Barrow, J.D. (1989) *Nuclear Physics B*, **323**, 242. https://doi.org/10.1016/0550-3213(89)90596-8

[6] Carroll, S.M., De Felice, A., Duvvuri, V., Easson, D.A., Trodden, M. and Turner, M.S. (2005) *Physical Review D*, **71**, Article ID: 063513. https://doi.org/10.1103/PhysRevD.71.063513

[7] Ishak, M. and Moldenhauer, J. (2009) *JCAP*, **01**, 024. https://doi.org/10.1088/1475-7516/2009/01/024

[8] Mannheim, P.D. (2012) *Foundations of Physics*, **42**, 388. https://doi.org/10.1007/s10701-011-9608-6

[9] Gasperini, M. and Veneziano, G. (1993) *Astroparticle Physics*, **1**, 317. https://doi.org/10.1016/0927-6505(93)90017-8

[10] Nojiri, S., Odintsov, S.D. and Sami, M. (2006) *Physical Review D*, **74**, Article ID: 046004. https://doi.org/10.1103/PhysRevD.74.046004

[11] Kanti, P., Mavromatos, N.E., Rizos, J., Tamvakis, K. and Winstanley, E. (1996) *Physical Review D*, **54**, 5049. https://doi.org/10.1103/PhysRevD.54.5049

[12] Cartier, C., Copeland, E.J. and Madden, R. (2000) *JHEP*, **01**, 035. https://doi.org/10.1088/1126-6708/2000/01/035

[13] Maeda, K.-I. and Ohta, N. (2004) *Physics Letters B*, **597**, 400. https://doi.org/10.1016/j.physletb.2004.07.038

[14] De Felice, A. and Hindmarsh, M. (2007) *JCAP*, **06**, 028. https://doi.org/10.1088/1475-7516/2007/06/028

[15] Li, B., Barrow, J.D. and Mota, D.F. (2007) *Physical Review D*, **76**, Article ID: 044027. https://doi.org/10.1103/PhysRevD.76.044027

[16] Zhou, S.-Y., Copeland, E.J. and Saffin, P.M. (2009) *JCAP*, **07**, 009. https://doi.org/10.1088/1475-7516/2009/07/009

[17] Buchdahl, H. (1964) *Acta Mathematica*, **85**, 63. https://doi.org/10.1007/BF02395741

[18] Schmidt, H. (1990) *Classical and Quantum Gravity*, **7**, 1023. https://doi.org/10.1088/0264-9381/7/6/011

[19] Horava, P. (2009) *Physical Review D*, **79**, Article ID: 084008. https://doi.org/10.1103/PhysRevD.79.084008

[20] Mukohyama, S. (2010) *Classical and Quantum Gravity*, **27**, Article ID: 223101. https://doi.org/10.1088/0264-9381/27/22/223101

[21] Bergmann, P.G. (1968) *International Journal of Theoretical Physics*, **1**, 25. https://doi.org/10.1007/BF00668828

[22] Nordtvedt, K. (1970) *The Astrophysical Journal*, **161**, 1059. https://doi.org/10.1086/150607

[23] Wagoner, R.V. (1970) *Physical Review D*, **1**, 3209. https://doi.org/10.1103/PhysRevD.1.3209

[24] Bailin, D. and Love, A. (1987) *Reports on Progress in Physics*, **50**, 1087. https://doi.org/10.1088/0034-4885/50/9/001

[25] Overduin, J.M. and Wesson, P.S. (1997) *Physics Reports*, **283**, 303. https://doi.org/10.1016/S0370-1573(96)00046-4

[26] Rubakov, V.A. (2001) *Physics-Uspekhi*, **44**, 871. https://doi.org/10.1070/PU2001v044n09ABEH001000

[27] Maartens, R. and Koyama, K. (2010) *Living Reviews in Relativity*, **13**, 5. https://doi.org/10.12942/lrr-2010-5

[28] Eisenhart, L.P. (1958) Non-Riemannian Geometry. American Mathematical Society, Providence.

[29] Ehlers, J. and Schild, A. (1973) *Communications in Mathematical Physics*, **32**, 119. https://doi.org/10.1007/BF01645651

[30] Will, C.M. and Nordtvedt, K. (1972) *The Astrophysical Journal*, **177**, 757. https://doi.org/10.1086/151754

[31] Will, C.M. (1979) General Relativity: An Einstein Centenary Survey. Cambridge University Press, Cambridge.

[32] Will, C.M. (1993) Theory and Experiment in Gravitational Physics. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511564246

Journal of Modern Physics, 2020, 11, 1031-1043
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Information-Based Numerical Distances between Equilibrium and Non-Equilibrium States

Angel Ricardo Plastino[1,2], Gustvo Luis Ferri[3], Mario Carlos Rocca[2,4,5], Angelo Plastino[2,4,6]

[1]CeBio y Departamento de Ciencias Básicas, Universidad Nacional del Noroeste de la Prov. de Buenos Aires, Junin, Argentina
[2]Argentina's National Research Council (CONICET), La Plata, Argentina
[3]Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa, Junin, Argentina
[4]Departamento de Física, Universidad Nacional de La Plata, La Plata, Argentina
[5]Departamento de Matemática, Universidad Nacional de La Plata, La Plata, Argentina
[6]SThAR - EPFL, Lausanne, Switzerland
Email: angeloplastino@gmail.com

How to cite this paper: Plastino, A.R., Ferri, G.L., Rocca, M.C. and Plastino, A. (2020) Information-Based Numerical Distances between Equilibrium and Non-Equilibrium States. *Journal of Modern Physics*, **11**, 1031-1043.
https://doi.org/10.4236/jmp.2020.117065

Received: June 20, 2020
Accepted: July 11, 2020
Published: July 14, 2020

Abstract

We consider a typical master equation describing thermal time-evolution. In parallel, we also consider a quasi static canonical description of the same problem. We are able to devise a way of numerically comparing these two treatments and concoct a distance-measure between them. In this way, one is in a position to know how far or close equilibrium and off-equilibrium can get. The first, rather surprising observation, is that our systems lose structural details as N grows. Also, the time-evolution of the distance between the two pertinent probability distributions is quite sensitive to the heating-cooling process.

Keywords

Information Theory, Master Equation, Thermal Time-Evolution, Equilibrium-Off Equilibrium Distances

1. Introduction

1.1. Preliminaries

Researchers often appeal to master equations (ME) to obtain an equation of motion for the reduced density operator. Or for the probability distribution (PD) of a subsystem of interest A in interaction with (a usually much larger) subsystem B (heath bath, for instance). The key issue is that our system of interest A is in a

situation of off-equilibrium. The literature on the subject is really enormous (Google Scholar returns lists around 30,000 links). Thus, we content ourselves with citing [1] and references therein. The issue at hand is how to extract relevant information on system A from the pertinent von Neumann equation. Consequently, our aim is to derive the time evolution of the PD for the entire system A + B, in a way which guarantees that normalization (amongst other properties) is preserved at any time *t*. The solution to this problem is found in the so-called master equation (ME) technique [1]. A popular, but not rigorous ME-approach can be used in the case of physical situations for which interacting systems A and B are known and well-defined so that one constructs the corresponding ME-equation of motion for the PD [1]. A beautiful instantiation of the ME-procedure is presented by Takada, Conradt, and Richet (TCR) in [2], for example, we will follow here without further ado.

1.2. TCR Main Ideas

TCR consider a two-level (1 and 2) model (system A) in contact with a reservoir B of temperature T. If the population of the excited state is computed, then that of the ground state becomes automatically fixed. The transition rate is the crucial parameter governing the degree of non-equilibrium. Denote by p_1 and p_2 the concomitant occupation probabilities of the lower and upper wells, respectively.

TCR imagine a heating and cooling process in which the reservoir's temperature is a function of time. T first grows, attains a maximum value, and then decreases. One can interpret this scenario as that of a particle moving in an asymmetric double-well potential. Site 1 is the bottom of the first well, whose energy is E_1. Likewise, site 2 is the bottom of the second well, at a higher energy E_2. Then, E_1 is the potential energy barrier to be overcome between states 1 and 2. System A subsequently evolves with a concomitant energy decrease to E_2, leading to state 2. See **Figure 1**. TCR write the associated master equation as

$$\mathrm{d}p_1/\mathrm{d}t = -a_1 p_1 + a_2 p_2;\ p_1 + p_2 = 1, \tag{1}$$

with

$$a_i = \exp\left[-E_i/k_b T\right],\ i = 1, 2,\ k_B = \text{Boltzmann's constant}. \tag{2}$$

1.3. Present Goal

Inspired by [2], we wish to address here a different problem. We will tackle a quantum many-body system of interacting fermions, advanced in [3], for which the interaction is ruled by an SU2 algebra. The system is heated and cooled as depicted in **Figure 1**, and we will deal with a master coupled system of N equations (not just two as in [2]). The thermal process will compete with the fermion-fermion interaction's effects. A canonical ensemble treatment of this model is reported in [4]. In parallel, for every different temperature T, we consider a fictitious coupled system-reservoir in thermal equilibrium at such T (quasi-static approximation (qsa). We will be able to devise, via Information Theory quantifiers, several notions of distance between the "master equation probability distribution"

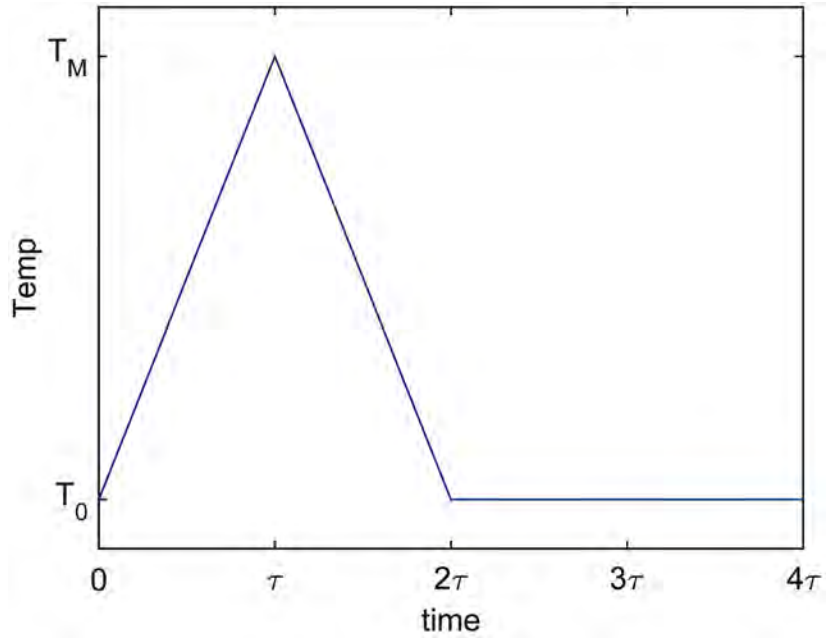


Figure 1. Time evolution of the heath-bath temperature $T_0 < T < T_M$ between a time $0 < t < 2\tau$. $T = T_0$ for $T < 2\tau$.

(MEPD) and the "qsa" PD, which of course, will yield numerical values that measure how far or close the two treatments are.

Epistemology tells us that classification is an essential feature of scientific endeavor [5]. For such a task we need numbers, and the distances that we are looking for provide such numbers, perhaps for the first time ever in the present context.

1.4. Paper's Organization

Sections 2 and 3 describe appropriate details of our exactly solvable quantum many-body system, that will serve as a laboratory to test our ideas of thermal distance between off equilibrium and equilibrium distances (OEED). Section 4 illustrates about the notion of statistical complexity. Section 6 is of the essence, as it presents our main ideas: that of a) a master equation and b) of quantifiers of OEED. The main results are exhibited and discussed in Section 7 and, finally, some conclusions are drawn in Section 8.

2. *M*-Fermions' Exactly Solvable Model

The model advanced in Ref. [3] and further discussed in Ref. [4], considers M fermions distributed amongst ($2M$)-fold degenerate single-particle (sp) levels, separated by a sp energy gap ϵ. Two quantum numbers μ and p are attached to a generic single particle state. The first adopts the values $\mu = -1$ (lower level) and $\mu = +1$ (upper level). p, usually called quasi-spin or pseudo spin, singles out a state within the M-fold degeneracy. The couple p, μ can also be viewed as a "site" that is either occupied or empty. One has

$$M = 2J, \tag{3}$$

where J denotes an "angular momentum". One introduces now the quasi-spin operators

$$\hat{J}_+ = \sum_p C^\dagger_{p,+} C_{p,-}, \tag{4}$$

$$\hat{J}_- = \sum_p C^\dagger_{p,-} C_{p,+}, \tag{5}$$

$$\hat{J}_z = \sum_{p,\mu} \mu C^\dagger_{p,\mu} C_{p,\mu}, \tag{6}$$

$$\hat{J}^2 = \hat{J}_z^2 + \frac{1}{2}\left(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+\right), \tag{7}$$

the eigenvalues of $\hat{J}^2$ being of the form $J(J+1)$. The Hamiltonian of reference [3], a spin-flip one, reads

$$\hat{H} = \epsilon \hat{J}_z - V_s\left(\frac{1}{2}\left(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+\right) - \hat{J}\right), \tag{8}$$

or, with $V = V_s/\epsilon$ (equivalently, take $\epsilon = 1$),

$$\hat{H} = \hat{J}_z - V\left(\frac{1}{2}\left(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+\right) - \hat{J}\right), \tag{9}$$

so that the unperturbed ground state (ugs) ($V = 0$) becomes, according to Equation (3),

$$\left|J, J_z\right\rangle = \left|J, -M/2\right\rangle, \tag{10}$$

whose energy is

$$E_o = -M/2. \tag{11}$$

Doubly occupied p-sites are not allowed for. The Hamiltonian commutes with $\hat{J}^2$ and $\hat{J}_z$. This entails that the exact solution will belong to the J-multiplet of the unperturbed ground state. The multiplet's states are denoted as $\left|J, m\right\rangle$, and one of them will minimize the energy. The associated m value of this state will depend upon the coupling strength V of the interaction. As we just mentioned, for the ugs one has $m = -J = -M/2$. Evidently, the interaction-operator $\left(\hat{J}_+\hat{J}_- + \hat{J}_-\hat{J}_+\right)$ is a quasi-spin flipping operator. Thus, this operator becomes the more "effective" the more balanced the populations of the two-levels become.

T = 0-Phase Transitions

As V grows from zero, the ugs energy E_o is not immediately affected. It conserves its value till a critical V-specific value is attained, of $1/(M-1)$. At this juncture, the interacting gs suddenly turns out to be $\left|J, -M/2+1\right\rangle$. If V continues its growth, new phase transitions (PT) take place. The PT between $J_z = -k$ and $J_z = -k+1$ ensues at $V = 1/(2k-1)$. The successive PT's processes ends up when we attain either $J_z = 0$ ($V_{crit} = 1$ for integer J), or $J_z = -1/2$ ($V_{crit} = 1/2$ for odd J). Thus, at such juncture one has, independently of J [3]:

$$V_{crit} = 1/2 \text{ for half-integer } J \text{ or } V_{crit} = 1 \text{ for integer } J. \tag{12}$$

3. Model's Treatment at Finite Temperatures *T*

To repeat: double occupancy of a *p*-site is not permitted. Thus, the Hamiltonian matrix' size is $(2J+1)\times(2J+1)$. The only way to get different *J*'s is to have double occupancy [4]. Following this last reference, the $J = N/2$ multiplet is the only one we need to consider.

The free energy *F* and the partition function *Z*, with β the inverse temperature, and $k_B = 1$ (Boltzmann) are given by

$$F = -T \ln Z = -T \ln Trace\left(\exp\left(-\beta \hat{H}\right)\right). \tag{13}$$

In the Trace we sum over the J_z quantum number *m*. As *H* commutes with both *J* and J_z one finds

$$Z = \sum_{m=-J}^{m=J} \exp\left(-\beta E_m\right), \tag{14}$$

where the energy E_m are

$$E_m = m - V\left(J\left(J+1\right) - m^2 - J\right). \tag{15}$$

Our all important pertinent probabilities P_m are [4]

$$P_m = \frac{\exp\left(-\beta E_m\right)}{Z}, \tag{16}$$

for all $m = -J, -J+1, \cdots, J-1, J$. The entropy *S* is

$$S = -\sum_{m=-J}^{m=J} P_m \ln P_m. \tag{17}$$

4. Meaning of the Statistical Complexity Measure

Sometimes one wishes to grab hold of a system's correlation structures just as entropy grasps disorder. Why? Because such correlations strongly influence the main features of the prevailing PD describing physical processes. It is obvious that the opposite extremes of perfect order and maximal randomness do not manifest notable structural correlations [6]. In between these two instances, a wide range of structural degrees (SD) usually exist, that are in turn reflected by the traits of the prevailing PD *P*.

The authors of Ref. [6], invented a quite adequate functional $F[P]$ that can apprehend correlations just as Shannon's entropy encapsulates randomness, which indeed constituted an important breakthrough. Their ideas were conceptualized via the definition of López-Ruiz, Mancini, and Calbet (LMC) [6] of what became called the statistical complexity *C*.

LMC *C* individualized and quantified the respective contributions of entropy and structure. The last one was described by a quantity called disequilibrium *D*. Their concept of statistical complexity *C* was widely accepted (for a sample see, for instance, Refs. [6]-[27]). *C* vanishes in the two situations of perfect order and maximum disorder, being defined as the product of Shannon's entropy *S* and the

disequilibrium *D*. More specifically, the latter is a measure in probability space of the distance measure the prevailing PD and the uniform PD, so that one writes

$$C = SD, \tag{18}$$

with (see (17)), in our case, the uniform probabilities are $P^{(u)} = 1/(2J+1)$ for all m between $-J$ and J, so that, the LMC disequilibrium is

$$D = \sum_{m=-J}^{m=J} \left(P_m - P^{(u)}\right)^2, \tag{19}$$

while

$$S = -\sum_{m=-J}^{m=J} P_m \ln P_m. \tag{20}$$

For details, properties, and others applications of *C*, see Refs. [6].

5. Master Equations for *N*-Levels Systems

We consider that our system is at the equilibrium temperature T_0 at $t = 0$. Then he system is heated and reaches a temperature T_M at $t = \tau$. At this stage, the systems cools-off, reaching a temperature $T_f = T_0$ at time $t = 2\tau$. This is the temperature of the heat-reservoir at all later times.

Inspired by the 2-level systems treatment of [2], we tackle now a system of N levels $l_i\,(i = 0,1,2,\cdots)$, with level-energies $E_0, E_!$, etc., and selections rules that allow for only certain kinds of transitions. In the usual parlance of quantum many body theories, we permit only transitions from a state with k-particle-holes (p-h) to ones with either $k+1$ or $k-1$ p-h, so that

$$\frac{\mathrm{d}p_n}{\mathrm{d}t} = p_{n-1}\exp\left(-\beta E_n\right) - p_n \exp\left(-\beta E_{n-1}\right), \tag{21}$$

and we follow in this way till we face, for the last 3 steps,

$$\frac{\mathrm{d}p_2}{\mathrm{d}t} = p_1 \exp\left(-\beta E_2\right) - p_2 \exp\left(-\beta E_1\right),$$

$$\frac{\mathrm{d}p_1}{\mathrm{d}t} = p_0 \exp\left(-\beta E_1\right) - p_1 \exp\left(-\beta E_0\right),$$

$$\frac{\mathrm{d}p_0}{\mathrm{d}t} = -\left[p_{n-1}\exp\left(-\beta E_n\right) - p_n \exp\left(-\beta E_{n-1}\right)\right] - \cdots$$

$$-\left[p_0 \exp\left(-\beta E_1\right) - p_1 \exp\left(-\beta E_0\right)\right].$$

The last equation guarantees normalization of probabilities.

The initial conditions are $p_i(0) = \exp\left(-E_i/kT\right)/Z$ for all *i*.

β depends upon time, as described at the beginning of this Section and we work with

$$\beta(t) = 1/T(t),$$

and

$$E_j = j; \text{ all } j.$$

If β were constant, then the time-dependent probabilities $p_i(t)$ would relax to stationary Gibbs-distributions.

The numerical calculations consider the cases of $N=3,4,5,6$, corresponding respectively to $J=1,3/2,2,5/2$ for our model above.

Distance Quantifiers

We deal with the results of our master equation (ME) above, that are to be compared to the quasi-static (st) results that arise out of considering always Boltzmann-Gibbs equilibrium-situations at the temperature $T(t)$ for all *t*. As distance quantifiers we will employ

- The probabilities' differences 1) Global $D_P=\sum_m\left[P_m^{ME}(t)-P_m^{st}(t)\right]^2\Big/\left(P_m^{ME}(t)\right)^2$, and 2) Individual $d_P^m=\left[P_m^{ME}(t)-P_m^{st}(t)\right]^2\Big/\left(P_m^{ME}(t)\right)^2$,
- the entropy *S*,
- the free energy *F*,
- the mean energy *U*, and
- the statistical complexity *C*.

If we generically call Q to any of these quantifiers, the distances are of the form

$$d_Q(t)=\frac{\left[Q_{ME}(t)-Q_{st}(t)\right]^2}{Q_{ME}(t)^2}, \tag{22}$$

$$D_Q=\int_0^{20\tau}\mathrm{d}t\,d_Q(t), \tag{23}$$

where, obviously, the sub-index ME (or just "M") refers to master equation's results, and the sub-index "st" (or just "s") to quasi-static ones.

We expect the distances to be sensitive to

1) Changes in the behavior of *T* with *t* and

2) Structural system's changes with the coupling constant *V*,

3) The speed of the heating-up and cooling-off process, regulated by the parameter τ. The shorter τ, the faster the speed.

6. Distances' Results

Note that, with reference to our model above, we have $N=2J+1$. **Figure 2** refers to $N=3$ and the individual probabilities-distance D_P^m vs. *t*. It is clearly seen that there is much sensitivity to the details of the heating process. **Figure 3** refers also to $N=3$ but with the global D_P vs. *V*. Here the ensuing picture is of a more complex nature. There is a system's phase transition (PT) at $V=1$, and our distance tends to decrease as we approach the PT. Thus we see that D_P is sensitive to the internal dynamics of the system. This is no trivial issue. Why should the OE-equilibrium distance behave in such a manner? This is a new fact discovered here, as far as we know. But the picture becomes even more complex when we consider the speed of the heating process. If it is large enough, the distance OE-E grows again with *V* after the PT, but if it is low enough, it continues

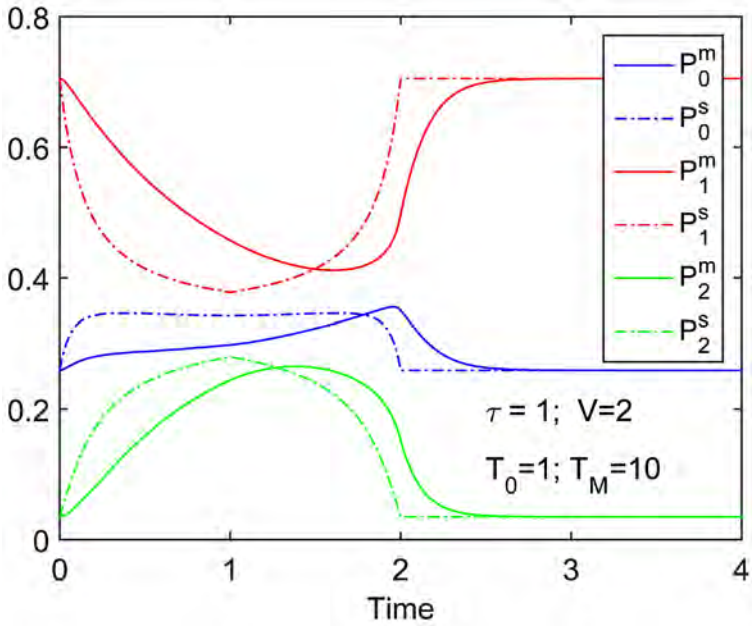


Figure 2. The distance D_{OE_E} (see text) off equilibrium-equilibrium for different τ values.

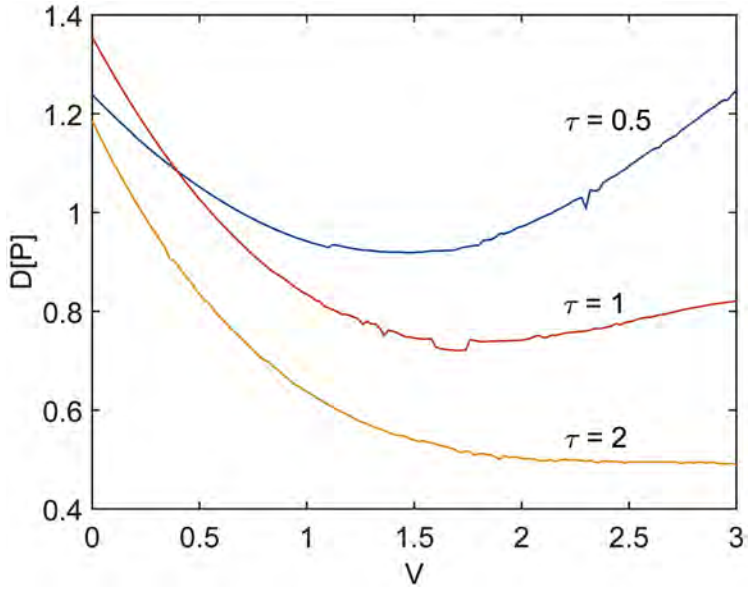


Figure 3. $N = 3$: the distance D_{OE_E} vs. the coupling constant V for several τ values (see text).

decreasing as V grows. *We conclude that the distance quantifiers is sensitive both to the system's internal dynamics and to the details of the heating process.*

Starting now with $N = 4$ we pass to watch in **Figure 4** the performance with time of the other quantifiers, of thermal origin, namely *S*, *F*, *U*, and *C*. As compared with the probabilities-distance, the thermal-distance (TD) is much more sensitive to the details of the heating process. When it finishes, at $t = 2\tau$, the four different TD vanish. We pass now to **Figure 5** so as to analyze, for $N = 3$,

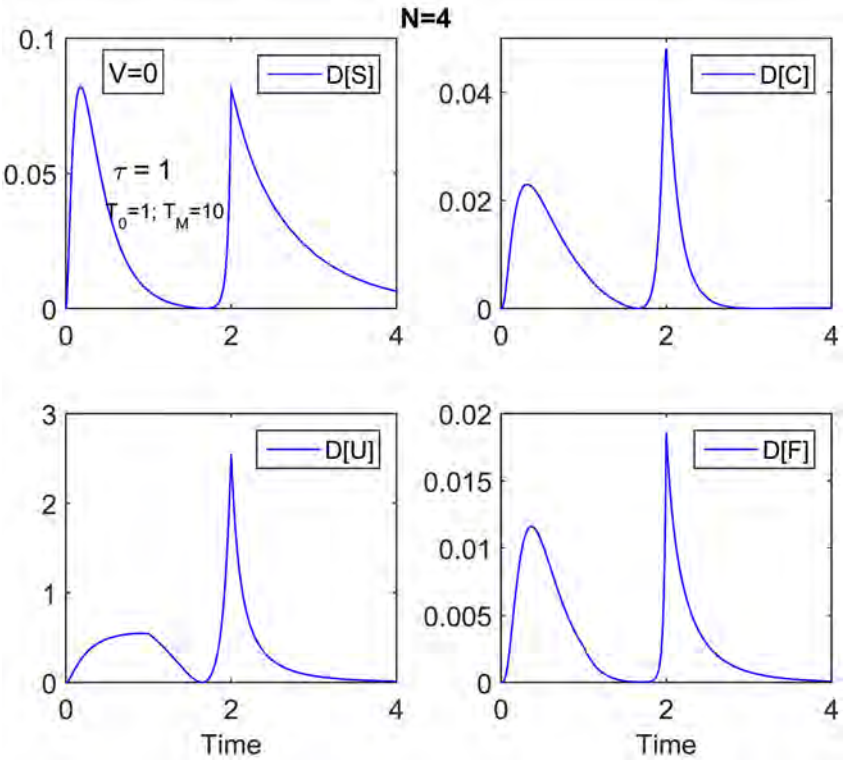


Figure 4. The four thermal distances vs. time for $N=4$ (see text).

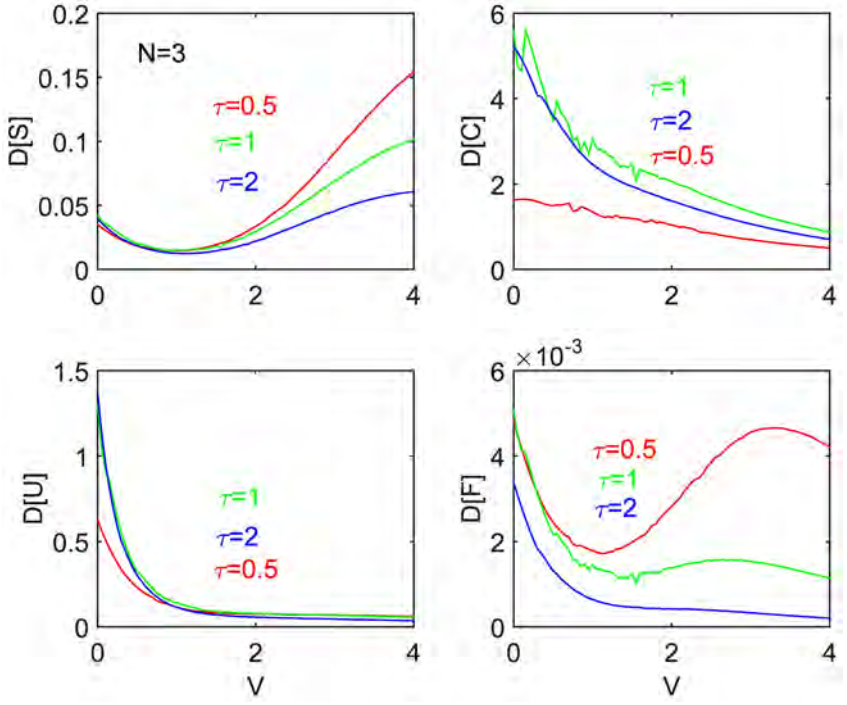


Figure 5. $N=3$: the four thermal distances vs. the coupling constant V for several τ values (see text).

the behavior of the four thermal distances. We see that each of them behaves in a distinctive way, that forces individual consideration.

- *S*-distance: it diminishes as V grows from zero till reaching the phase transition (PT). After wards it grows again as V continues increasing. At the PT the distance is minimal. The after-PT grows is the more pronounced the speedier the heating process (HP). Even after it has finished, S seems to keep memory of it, since its growth differentiates amongst the different τ's.
- *F*-distance: It senses the PT if the HT is fast. Otherwise, it does not. It clearly is able fo distinguish amongst the different τ's.
- *U*-distance: tends to vanish as V grows. It is rather indifferent both to the τ-value and to the PT.
- *C*-distance: this is supposed to be the most sophisticated statistical of our four tools. Here it only tells us that the difference between the off-equilibrium complexity and the equilibrium one tend to become identical as the coupling constant grows. In other words, if the system is tightly bound, whether one heats it up or not becomes less and less relevant as the bonding augments, which sounds reasonable.

Figure 6 is identical to Figure 5, except for the fact that N is larger and equals 5. There are two, not just one, PTs now, both for $V<1$, that do not seem to leave any trace in the graphs. The differences between what the different thermal indicators say is smaller here. All of them indicate that if the system is tightly bound, whether one heats it up or not becomes less and less relevant as the bonding augments, as we saw above for $N=3$, but only in the *C*-case. Here, once again, the statistical complexity appears to be the mpst sensitive of the indicators, as the *C*-distance is the only one of the four here used able to detect differences amongst the distinct τ values. These conclusions are reinforced if

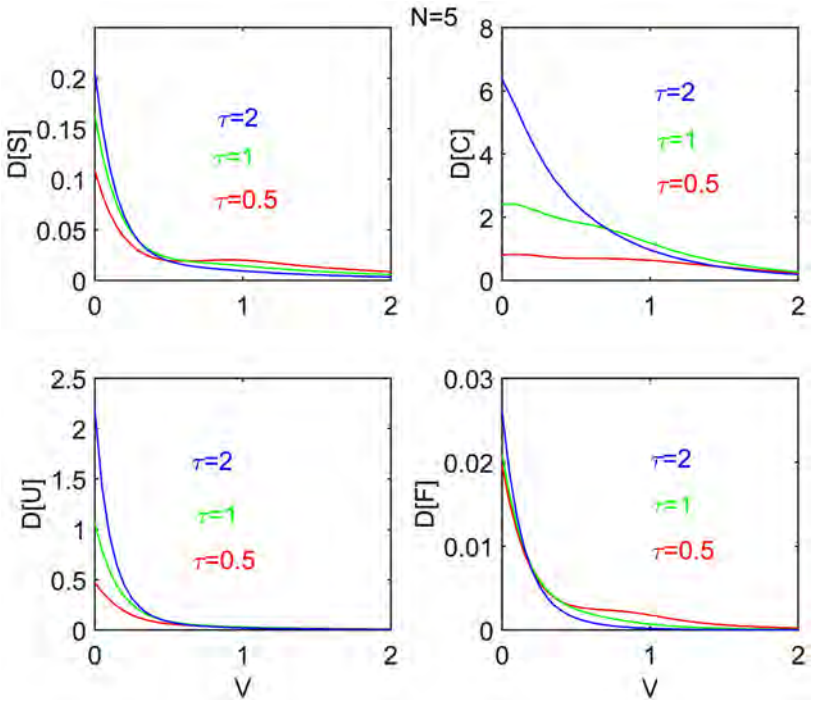


Figure 6. $N=5$: same details as in Figure 5, but for a larger N-value (see text).

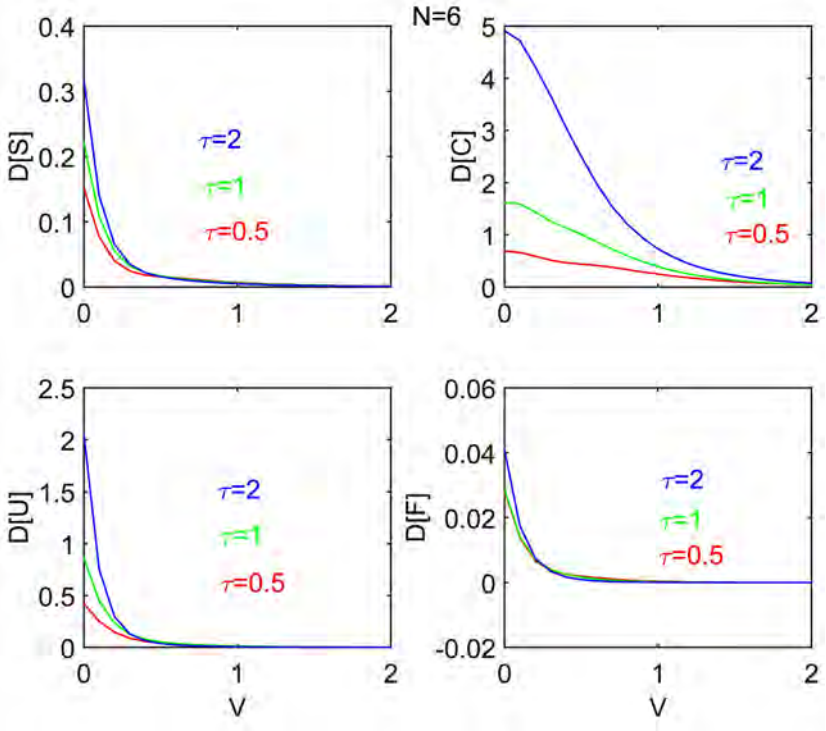


Figure 7. $N=6$: same details as in **Figure 5**, but for a larger $N=6$-value (see text).

we confront now the case $N=6$. Emphasize that C is undoubtedly the most sensitive quantifier (**Figure 7**).

7. Conclusions

- The first, rather surprising observation, is that our systems lose structural details as Ngrows.
- The time-evolution of the distance between the two pertinent probability distributions is quite sensitive to the heating-cooling process.
- The shorter the heating-coolong period, the more sensitive the probabilities-distance quantifier becomes to the internal systems' structure, as revealed by the ground-state phase transition.
- The four thermal distances time-evolutions are also quite sensitive to the heating-cooling process.
- The four thermal distances evolutions with a growing coupling constant are quite sensitive to the internal dynamics and to g the heating-cooling details for $N=3$.
- This sensitivity is gradually lost as Ngrows.
- For $N\geq 4$, the four thermal distances evolutions with a growing coupling constant rapidly vanish. The strongly interacting systems seem no to care whether it is heated or cooled.
- This entails that, the larger the coupling constant V, the more rapidly equilibrium is attained, as evidences by the diminution of the values of our thermal distance quantifiers. However, $C$is the most sensitive of the four indicators.

Conflicts of Interest

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

References

[1] Kryszewski, S. and Czechowska-Kryszk, J. (2008) Master Equation—Tutorial Approach.

[2] Takada, A., Conradt, R. and Richet, P. (2013) *Journal of Non-Crystalline Solids*, **360**, 13. https://doi.org/10.1016/j.jnoncrysol.2012.10.002

[3] Plastino, A. and Moszkowski, S.A. (1978) *Il Nuovo Cimento*, **47**, 470. https://doi.org/10.1007/BF02896236

[4] Pennini, F. and Plastino, A. (2018) *Physica A*, **491**, 305. https://doi.org/10.1016/j.physa.2017.09.041

[5] Ziman, J. (2002) Real Science. Cambridge Univ. Press, Cambridge.

[6] López-Ruiz, F.R., Mancini, H. and Calbet, X. (1995) *Physics Letters A*, **209**, 321. https://doi.org/10.1016/0375-9601(95)00867-5

[7] Feldman, D.P. and Crutchfield, J.P. (1998) *Physics Letters A*, **238**, 244. https://doi.org/10.1016/S0375-9601(97)00855-4

[8] Martin, M.T., Plastino, A. and Rosso, O.A. (2003) *Physics Letters A*, **311**, 126. https://doi.org/10.1016/S0375-9601(03)00491-2

[9] Kowalski, A., Martin, M.T., Plastino, A., Proto, A. and Rosso, O.A. (2003) *Physics Letters A*, **311**, 180. https://doi.org/10.1016/S0375-9601(03)00470-5

[10] Ribeiro, H.V., Zunino, L., Lenzi, E.K., Santoro, P.A. and Mendes, R.S. (2012) *PLoS ONE*, **7**, e40689. https://doi.org/10.1371/journal.pone.0040689

[11] Ribeiro, H.V., Zunino, L., Mendes, R.S. and Lenzi, E.K. (2012) *Physica A*, **391**, 2421. https://doi.org/10.1016/j.physa.2011.12.009

[12] Manzano, D. (2012) *Physica A*, **391**, 6238. https://doi.org/10.1016/j.physa.2012.06.058

[13] Borgoo, A., Geerlings, P. and Sen, K.D. (2011) *Physics Letters A*, **375**, 3829. https://doi.org/10.1016/j.physleta.2011.09.031

[14] Dehesa, J.S., López-Rosa, S.S. and Manzano, D. (2009) *The European Physical Journal D*, **55**, 539. https://doi.org/10.1140/epjd/e2009-00251-1

[15] Esquivel, R.O., Molina-Espritu, M., Angulo, J.C., Antolín, J., Flores-Gallegos, N. and Dehesa, J.S. (2011) *Molecular Physics*, **109**, 2353. https://doi.org/10.1080/00268976.2011.607780

[16] López-Ruiz, R. (2013) Concepts and Recent Advances in Generalized Information Measures and Statistics. Bentham Science Publishers, Sharjah, 147-168. https://doi.org/10.2174/978160805760311301012

[17] Sen, K.D. (2020) *Quantum Reports*, **20**, 313.

[18] Mitchell, M. (2009) Complexity: A Guided Tour. Oxford University Press, Oxford.

[19] Martin, M.T., Plastino, A. and Rosso, O.A. (2006) *Physica A*, **368**, 439. https://doi.org/10.1016/j.physa.2005.11.053

[20] Pennini, F. and Plastino, A. (2017) *Physics Letters A*, **381**, 212. https://doi.org/10.1016/j.physleta.2016.11.023

[21] Pennini, F. and Plastino, A. (2018) *Physica A*, **506**, 828.

https://doi.org/10.1016/j.physa.2018.05.003

[22] Pennini, F. and Plastino, A. (2017) *Physica A*, **488**, 85. https://doi.org/10.1016/j.physa.2017.07.005

[23] Pennini, F. and Plastino, A. (2019) *Entropy*, **21**, 558. https://doi.org/10.3390/e21060558

[24] Branada, R., Pennini, F. and Plastino, A. (2018) *Physica A*, **511**, 18. https://doi.org/10.1016/j.physa.2018.07.037

[25] Pennini, F. and Plastino, A. (2018) *Physica*, **491**, 305. https://doi.org/10.1016/j.physa.2017.09.041

[26] Pennini, F. and Plastino, A. (2017) *Physics Letters A*, **381**, 3849. https://doi.org/10.1016/j.physleta.2017.10.025

[27] Anteneodo, C. and Plastino, A.R. (1996) *Physics Letters A*, **223**, 348. https://doi.org/10.1016/S0375-9601(96)00756-6

Journal of Modern Physics, 2020, 11, 1044-1057
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Characterization of a New DC-Glow Discharge Plasma Set-Up to Enhance the Electronic Circuits Performance

A. A. Talab[1*], Ashraf Yahia[2], M. A. Saudy[2], M. Elsayed[2]

[1]Plasma Physics & Nuclear Fusion Department, Cyclotron Facility, NRC, Atomic Energy Authority, Cairo, Egypt
[2]Physics Department, Faculty of Science, Ain Sham University, Cairo, Egypt
Email: *azza_talab@yahoo.com

How to cite this paper: Talab, A.A., Yahia, A., Saudy, M.A. and Elsayed, M. (2020) Characterization of a New DC-Glow Discharge Plasma Set-Up to Enhance the Electronic Circuits Performance. *Journal of Modern Physics*, **11**, 1044-1057.
https://doi.org/10.4236/jmp.2020.117066

Received: May 12, 2020
Accepted: July 13, 2020
Published: July 16, 2020

Abstract

The (DC-GDPAU) is a DC glow discharge plasma experiment that was designed, established, and operated in the Physics Department at Ain Shams University (Egypt). The aim of this experiment is to study and improve some properties of a printed circuit board (PCB) by exposing it to the plasma. The device consists of cylindrical discharge chamber with movable parallel circular copper electrodes (cathode and anode) fixed inside it. The distance between them is 12 cm. This plasma experiment works in a low-pressure range (0.15 - 0.70 Torr) for Ar gas with a maximum DC power supply of 200 W. The Paschen curves and electrical plasma parameters (current, volt, power, resistance) characterized to the plasma have been measured and calculated at each cm between the two electrodes. Besides, the electron temperature and ion density are obtained at different radial distances using a double Langmuir probe. The electron temperature (KT_e) was kept stable in range 6.58 to 10.44 eV; whereas the ion density (n_i) was in range from 0.91×10^{10} cm^{-3} to 1.79×10^{10} cm^{-3}. A digital optical microscope (800×) was employed to draw a comparison between the pre-and after effect of exposure to plasma on the shaping of the circuit layout. The experimental results show that the electrical conductivity increased after plasma exposure, also an improvement in the adhesion force in the Cu foil surface. A significant increase in the conductivity can be directly related to the position of the sample surfaces as well as to the time of exposure. This shows the importance of the obtained results in developing the PCBs manufacturing that uses in different microelectronics devices like those onboard of space vehicles.

Keywords

DC Glow Discharge, Paschen Curve, Cold Plasma Characteristics, Double

Electric Probe, Printed Circuit Board (PCB) Properties, Electronic Plasma Application

1. Introduction

The non-thermal plasma, which is also named cold or non-equilibrium plasma, is weakly ionized with electron temperatures of a little eV and colder ion temperatures [1] [2]. The glow discharge plasma has different luminous and dark zones between cathode and anode as shown in **Figure 1**. It is produced when a direct current from a DC power supply is applied between these two parallel electrodes to a low-pressure gas [3] [4]. The energy is continuously transferred from a high-voltage DC power source to the electrons originating from the cathode that accelerates absorbing energy from the field, ionizing, exciting and undergoing elastic collisions with heavy particles and other electrons to produce the different zones. The glow discharge owes its name to the fact that plasma is luminous; the luminosity is produced because the electrons gain sufficient energy to generate visible light by excitation collisions which generate photons. Involved in the formation of glow discharges are the formation of restricted electric fields and plasma sheaths at each of the electrodes.

It is important after talking about the principals of producing the Dc glow discharge to refer to the Paschen's law. It is a function of gap length (d) between the two electrodes and the gas pressure (p) [*i.e.* $V_{bd} = f(Pd)$]. It describes the electric discharge between two conductive materials and determines the breakdown voltage (V_{bd}) at which the discharge process starts. Beside that (V_{bd}) depends on the density and type of the gas, the material of the two electrodes, the gap length between them, and the degree of preexisting ionization [5] [6] [7].

During this work, there are different plasma parameters that were measured and calculated. It was very important to begin with determining these parameters to know the suitable conditions of plasma to exposure printed circuit board (PCB) samples to it. An LCR meter (TEGAM) was used to measure the circuit resistance; a significant increase in the conductivity can be directly related to the position of the sample surfaces as well as to the time of exposure.

The double electric probe (DEP) is the most common probe type which has two identical isolated collectors with equal areas and draws no current from plasma, so it disturbs plasma only at its location [8] [9] [10]. From the slope of the IV characteristic curve of DEP as shown in **Figure 2**, kT_e can be obtained, and then n_i is evaluated by using the following equations [4] [11]:

$$\left.\frac{\mathrm{d}I}{\mathrm{d}V}\right|_{V=0} = \frac{e}{kT_e}\frac{i_{i1}i_{i2}}{i_{i1}+i_{i2}} \tag{1}$$

$$i_i = \frac{1}{2}en_iA\left(\frac{kT_e}{m_i}\right)^{0.5} \tag{2}$$

where (i_i) represents the saturation current; ($A = 2\pi rh$) is the area of the double

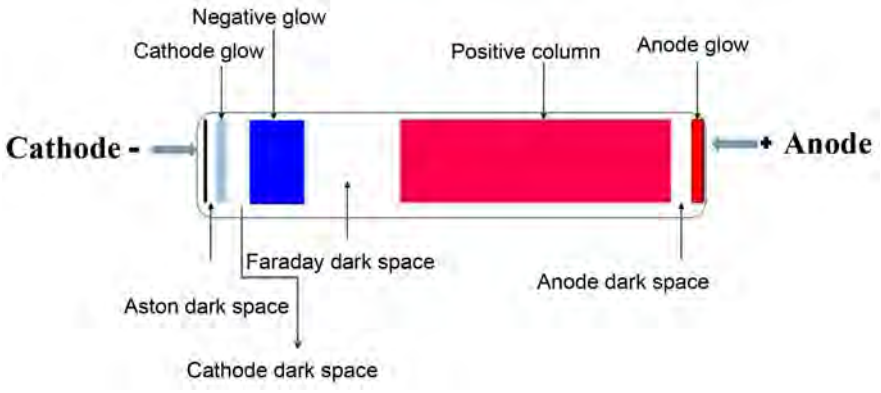


Figure 1. Regions of DC glow discharge plasma as in normal case. It is shown the arrangement of this regions and the area of each one.

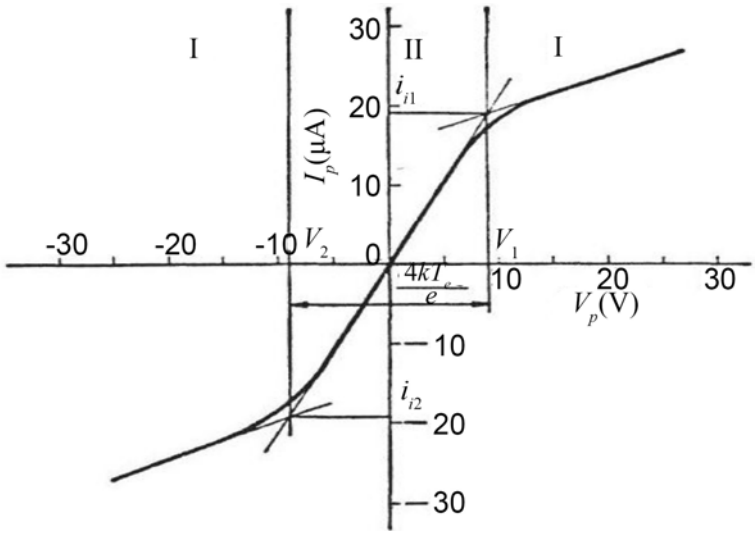


Figure 2. The typical IV characteristic curve of the double probe. It illustrates how KT_e is calculated. This curve was drowning from different waveforms which produced from experimental data.

Probe; r, h are the double probe's radius and length; and (m_i) is the ion mass of Ar [12].

Cold or low-pressure DC glow discharge plasma technology is applied widely in many fields, such as etching, surface treatment and activation, cleaning, electron beam source, sputtering spraying, adhesion improvement, etc. [13] [14] [15]. The adhesive forces bonding is a key joining technology in many manufacturing areas which include the automotive and aerospace industries, adding to biomedical and microelectronics applications [16]. The reactive species (electrons, atomic or molecular ions, atoms or molecules energetically excited) in the cold plasma can modify the superficial functional characteristics of the target material and promote the surface functionalization reactions by generating organic or inorganic thin layers because of the recombination of radicals or molecular fragments species on the surfaces of this material [17] [18] [19].

The printed circuit board (PCB) is a substrate that provides the necessary

electrical connections for an electronic circuit [20]. And the double layer PCB consists of an insulator layer between two conductive layers that are made of Cu foil which have excellent electrical and thermal conductivity properties. So, the surface treatment technique by plasma on Cu foil has become a useful tool to achieve a good adhesion for fine lines with a smooth surface. This is due to the rapid developments in high-frequency and high-speed transmission in electrical wire technology [21] [22] [23].

2. Experimental Setup

The DC glow discharge plasma experiment consists of three main parts, a discharge chamber, a vacuum system and an electric circuit. The cylindrical discharge chamber is made of stainless steel with an interior diameter 24.5 cm, a length 12.5 cm and a thickness 0.9 cm. It consists of two movable circular copper electrodes (cathode and anode) with a diameter 6.5 cm and a thickness 1 cm for each one. Where, the anode is left at a fixed location while the cathode moves along an axial length 12 cm.

The vacuum system consists of a rotary vacuum pump to evacuate the system, a needle valve to control the flow of gas, and an analog vacuum gauge controller unit to measure the vacuum and pressure of the gas inside the discharge chamber.

The electric circuit consists of a DC power supply 0 - 5 kV with a DC current 40 mA, a ballast resistor of 10 kΩ to limit the plasma current and a digital ammeter/voltmeter to measure the plasma current I_p and potential V_p. Adding to a potential divider consists of 50 kΩ and 22.3 MΩ resistors connected in series with each other. There is a high voltage probe, to convert the value of power supply voltage by ratio 1/1000, connects with a digital voltammeter to measure the V_{ps}.

The digital storage oscilloscope (DSO) is connected to a laptop using a USB interface to record the waveform of the breakdown voltage (V_{bd}) and the waveform of the probe current (I_{Probe}). The schematic diagram and setup of the plasma experiment are shown in **Figure 3**.

The DEP is made of two tungsten wires, each has 0.2 mm diameter with an equal length of 8 mm for its two tips and a 4 mm distance between them. The DEP circuit consists of a variable DC source which gives a probe voltage V_{Probe} = ±40 volts shown in **Figure 4(a)**. The DEP was placed vertically to the plasma and parallel to the electrodes inside the discharge chamber where it moves in a radial axis at distances (r = 2, 4, 6 and 8 cm) from the upper surface of the chamber as shown in **Figure 4(b)**. **Figure 4(c)** shows some examples of DEP waveforms at P = 0.3 Torr, d = 3 cm, r = 6 cm and I_p = 20 mA.

3. Results and Discussion

The characterization of the DC glow discharge plasma was determined by three parts, Paschen curves, and electrical plasma parameters, in addition to electron temperature and ion density of the plasma.

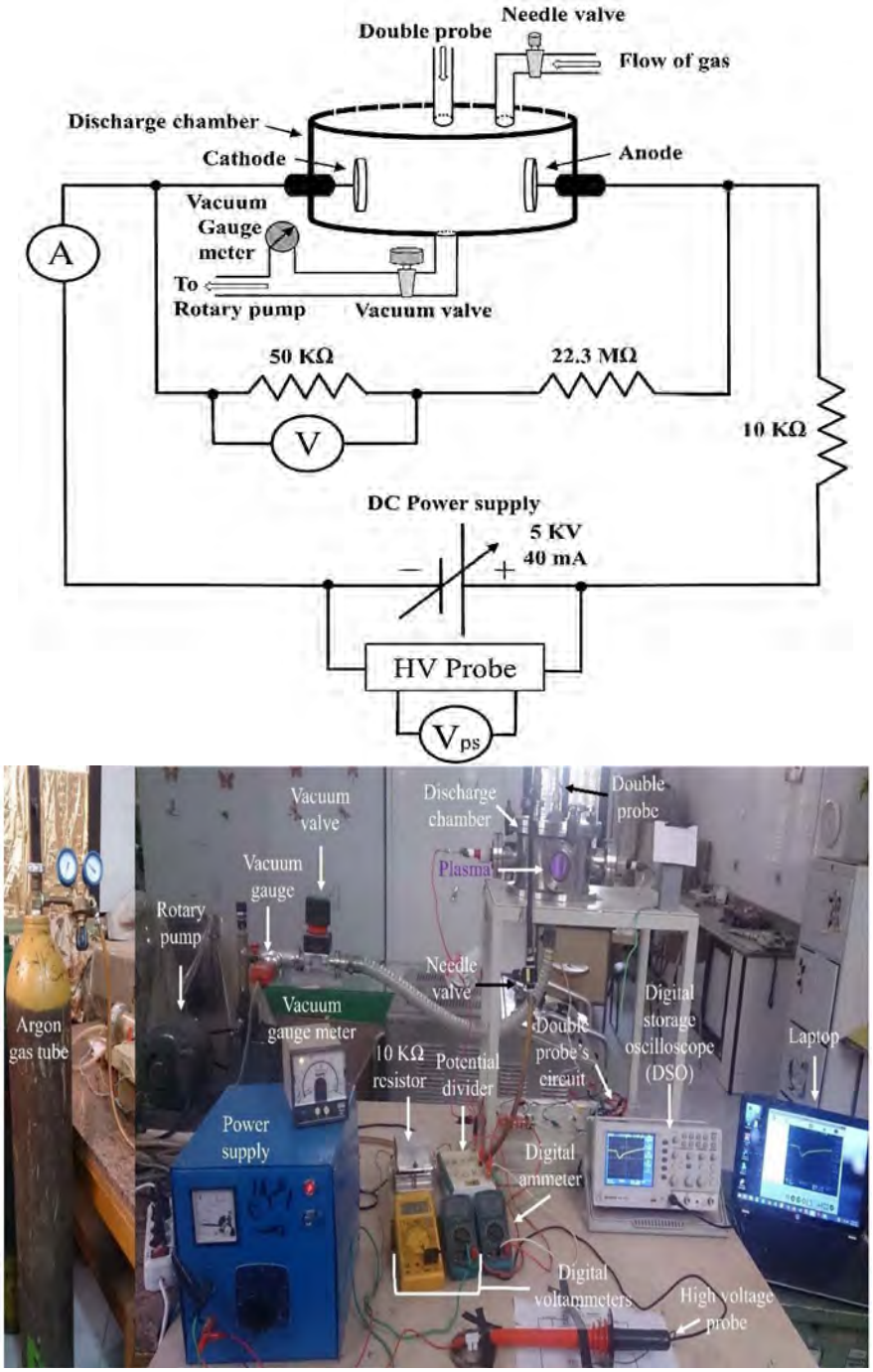


Figure 3. The schematic diagram and the photographic view of the plasma experiment.

The aim of this work is to establish a cold plasma experiment with specifications and characteristics that are in the normal range of previous DC glow discharge plasma experiments in other labs around the world. After validating our experiment, it can use to study the effect of exposing surfaces of different materials to this plasma. Here, the effect of the produced cold plasma on the electrical conductivity and the adhesive forces of a Cu foil were studied to improve the electronic circuits performance.

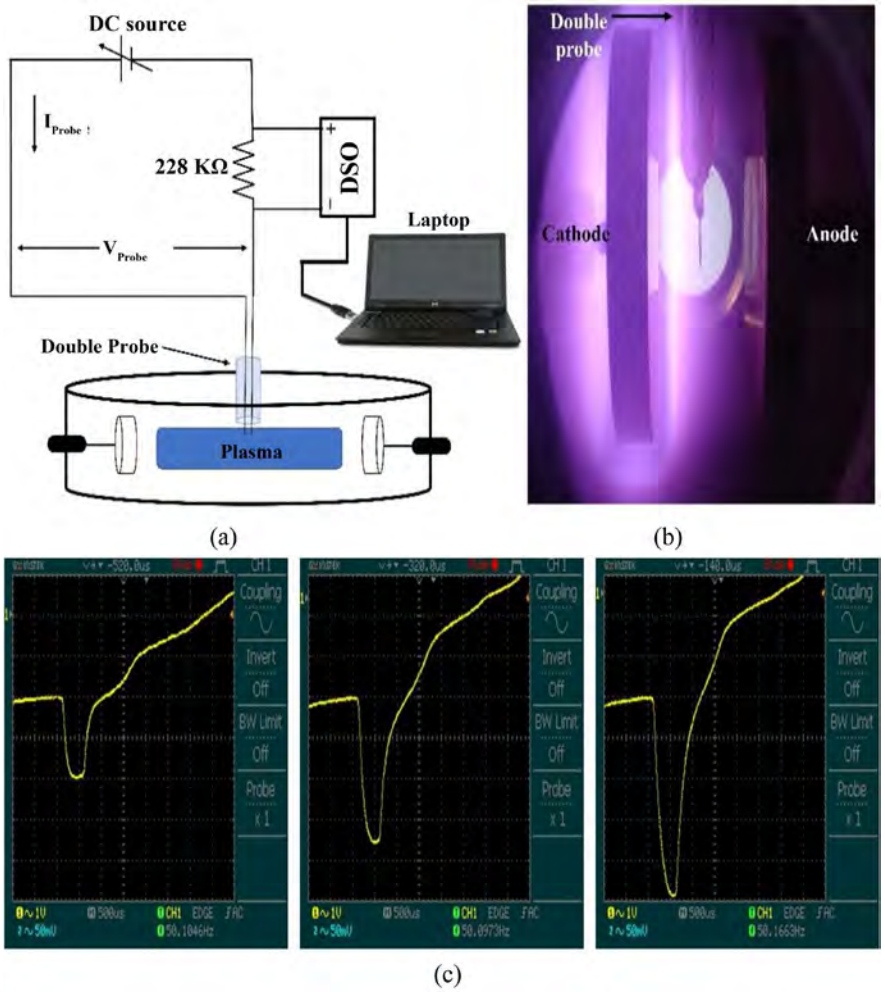


Figure 4. (a) DEP circuit, (b) DEP inside the plasma chamber, and (c) Examples of the waveform of the double electric probe.

3.1. Paschen Curves

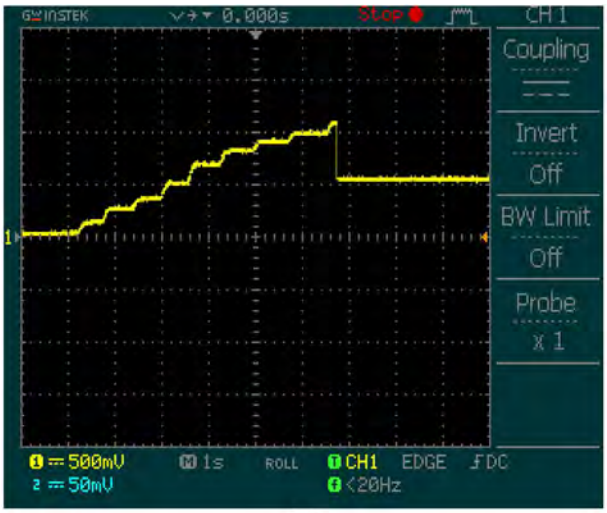
The Paschen curves were drawn at different gas pressures $P = 0.15$ to 0.70 Torr for each gap length $d = 3$ to 7 cm. Figure 5(a) shows the waveform of V_{bd} at $P = 0.70$ Torr and $d = 6$ cm, where, once the step of the voltage had occurred, the plasma was generated. Paschen curves at each gap length are shown in Figure 5(b), which cleared that the minimum breakdown voltage occurred at $P = 0.3$ Torr.

3.2. Electrical Plasma Parameters

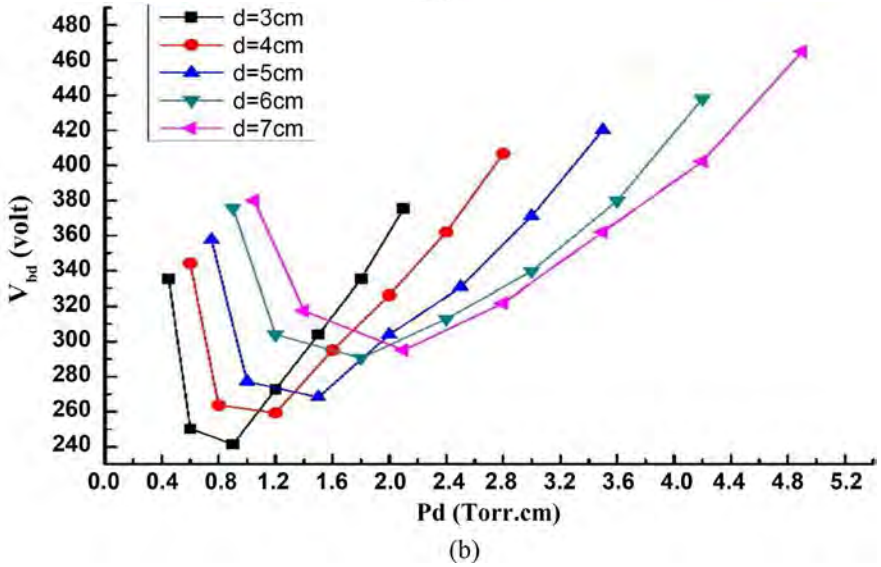
The electrical plasma parameters, namely, plasma voltage V_p, plasma current I_p, plasma resistance R_p and plasma power W_p were measured and calculated at each gap length with the operating conditions $P = 0.3$ Torr, and V_{ps} varies from 400 to 1800 volt.

The variation curves between the operating voltages V_{ps} with the electrical plasma parameters are shown in Figure 6. From these curves it obvious that:

1) Increasing V_{ps}, the number of ionized gas increased which resulted in

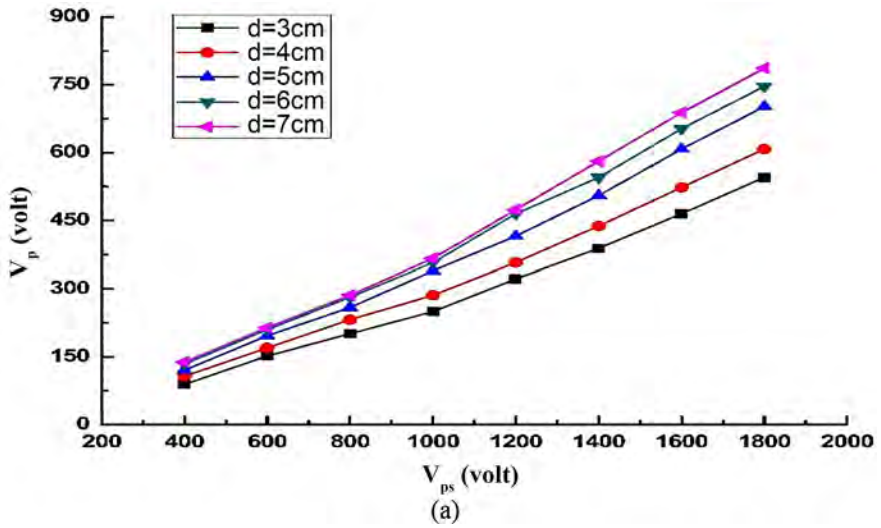


(a)



(b)

Figure 5. (a) Waveform of the breakdown voltage at Pd = 4.2 Torr.cm, and (b) Paschen curves at each gap length.



(a)

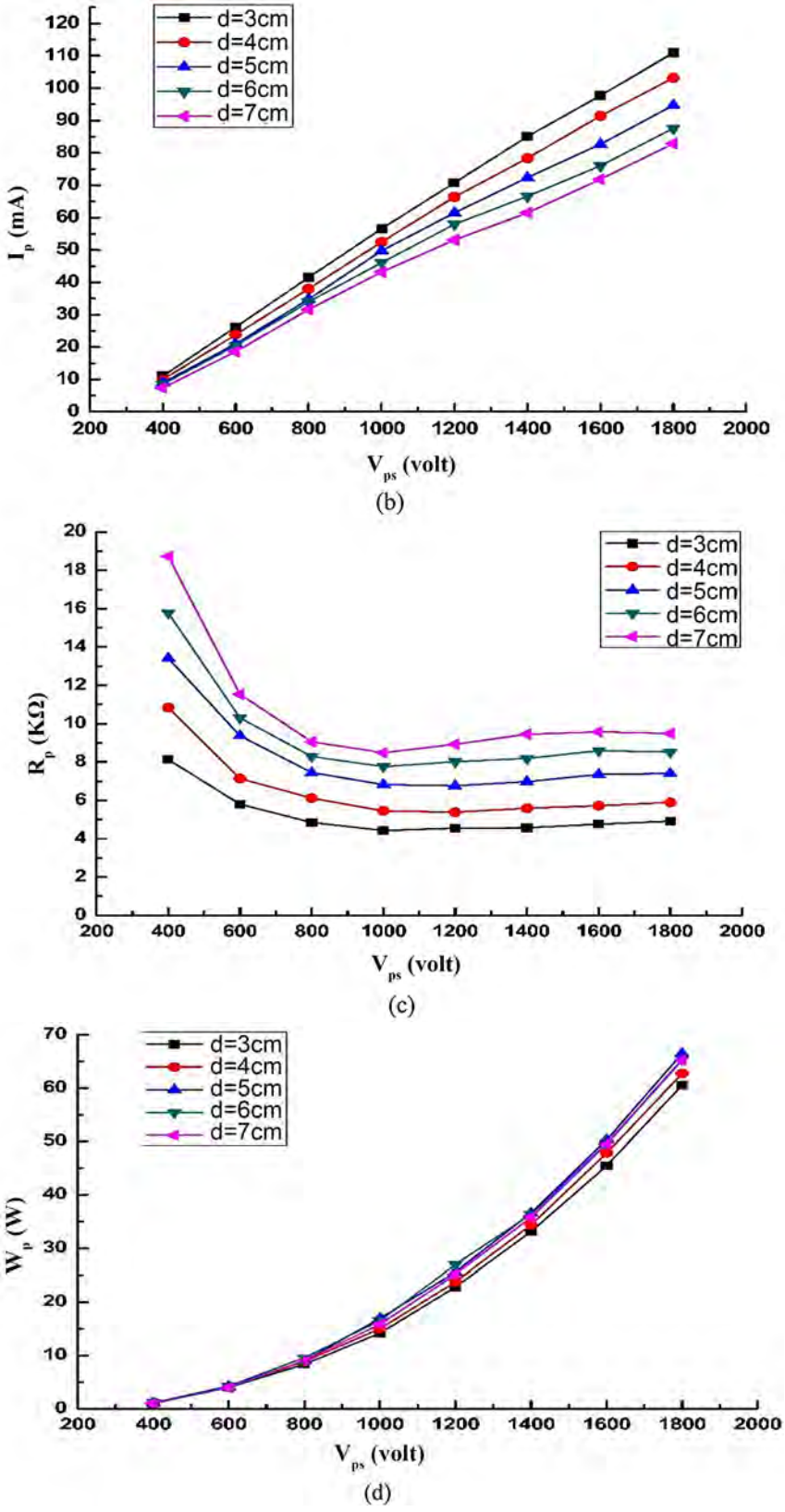


Figure 6. Shows at different gap lengths, the variation curves between the operating voltage V_{ps} and the electrical plasma parameters (a) the plasma current, (b) the plasma voltage, (c) the plasma resistance and (d) the plasma power.

increasing V_p, I_p and W_p. So, R_p is decreased with V_{ps}, because the resistance is inversely proportional to the current, until V_{ps} = 1000 volt, after that it was almost constant. This is because of the rate of variation V_p and I_p nearly constant.

2) For all below curves, there is almost no affecting change (d) in the region of variation V_{ps} from 400 to 1000 volt because the operating voltage in this region ionized the same number of Ar gas nearly. From 1000 volt the effect of this variation is clear because increasing the ionized numbers of Ar atoms leads to increasing the gap length.

3) From these results, the place where the samples exposed to the plasma radiation was defined (d = 5 cm), which is considered the center of the chamber.

3.3. Electron Temperature and Ion Density of Theplasma

The electron temperature (kT_e) and ion density (n_i) of plasma were appointed at various radial distances (r = 2 to 8 cm), for each gap lengths (d = 3 to 7 cm) with operating conditions (P = 0.3 Torr and I_p = 20 mA). The IV characteristic curves of the double electric probe are drawn in Figure 7 and the calculated values of the (kT_e) and (n_i) are tabulated in Table 1. The results show that the highest (n_i) and the lowest (kT_e) occurred at d = 5 cm where the samples were put to study the effect of plasma on them. Table 2 is a comparison between kT_e and n_i which were measured in our work versus that published before. This shows that the obtained results are validated with the previous work.

3.4. Effect of Plasma on Electronic Circuits Performance

A double layer PCB sample with thickness (1.55 mm) and area (15 cm^2) had two layers of Cu foil with thickness 3.5 × 10^{-2} mm of each one. It was placed inside the chamber at d = 5 cm (2 cm from the cathode) and exposed to the plasma,

Table 1. The different values of kT_e and n_i.

d (cm)	r (cm)	kT_e (eV)	$n_i \times 10^{10}$ (cm^{-3})
3	2	7.96	1.32
	4	8.29	1.14
	6	**10.44**	**0.91**
	8	9.34	1.01
5	2	6.58	1.58
	4	7.19	1.59
	6	**7.65**	**1.62**
	8	6.69	1.79
7	2	7.48	1.20
	4	7.67	1.32
	6	**7.99**	**1.26**
	8	7.33	1.24

Table 2. kT_e and n_i measured versus published.

	kT_e (eV)	n_i (cm^{-3})
Our calculated/measured	6.58 - 10.44	9.10×10^9 - 1.79×10^{10}
Kotp E F and Al-Ojeery A A (2012) [24]	1.90 - 3.50	6.40×10^9 - 7.40×10^9
Ghimire B *et al.* (2014) [25]	4.87 - 9.96	0.84×10^9 - 1.14×10^{11}
Nojiri K 2015 [26]	~1	10^9 - 10^{12}
Hashmi S *et al.* (2014) [27]	1.00 - 10.00	10^8 - 10^{13}
Oks E and Brown I (2002) [28]	7.50 - 8.50	10^{10}

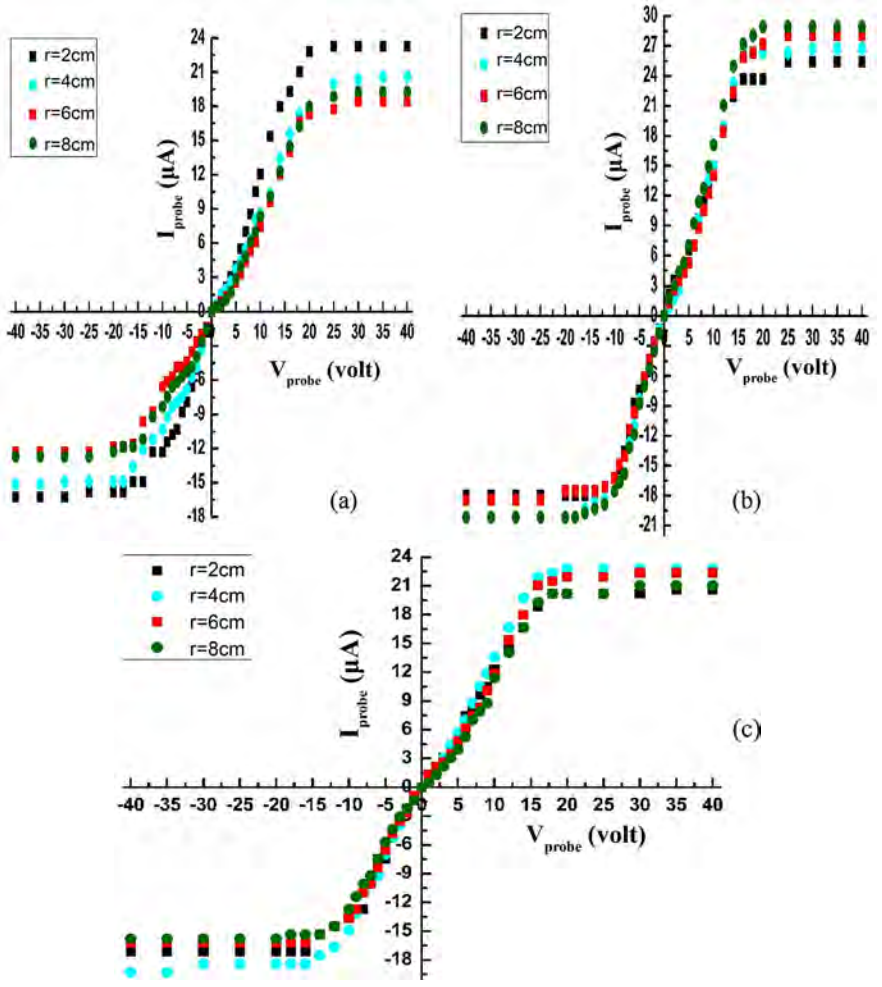


Figure 7. The IV characteristic curves at different gap lengths where (a) $d = 3$ cm, (b) $d = 5$ cm and (c) $d = 7$ cm.

with the operating conditions (V_{ps} = 1820 volt, V_p = 805 volt, I_p = 100 mA), for 10 min. **Figure 8** is a photographic view shows the difference in the effect of plasma on the two sample surfaces after exposure.

The electrical resistance (R) of Cu foil in a PCB sample was measured before and after exposure to the plasma by using the LCR meter type (TEGAM). The resistance of Cu foil before was 1.8×10^{-2} Ω while after exposure it was 0.7×10^{-2}

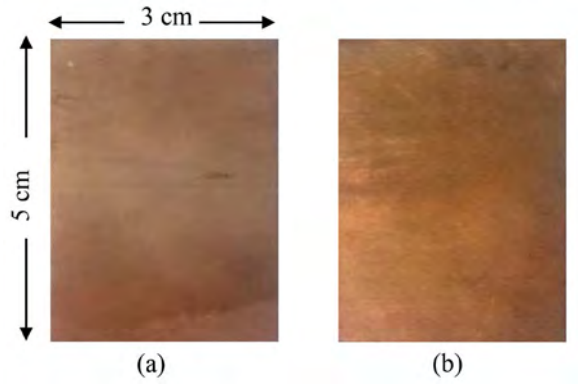


Figure 8. The photographic view of the two sides of PCB after exposed to the plasma. (a) Facing to the cathode; (b) Facing to the anode.

Ω for the layer in the front of the anode, however, the layer facing the cathode didn't have any change in its resistance. This means that the resistance of Cu foil after got exposed to the plasma is decreased, it allows greater current flow and the length/width (L/W) decreases [29] [30]. So, in turn, the electrical conductivity (σ) of the Cu foil is increased (by using the conductivity equation [31]) from 1.6×10^6 S/m to 4.1×10^6 S/m. So, the electronic components can be increased with the smaller size of Cu foil on a PCB. The reduction in the Cu foil size decreases the cost of PCB manufacturing which used in microelectronics applications (Integrated circuits, Cell phone, Laptop, etc.) [23].

There are many defects that may occur during printing the layout circuits on PCB. These defects are namely; breakouts, Pinhole, open-circuit, under etch, mouse bite, missing conductor, spurious copper, short-circuit, wrong side hole, conductors too close, missing hole and over etch [32]. Therefore, treating the PCB by exposure it to the plasma of specific characteristics is a necessary step to fix many of these problems.

Figure 9 shows the shape of the printing layout circuit on the Cu foil surfaces by using a digital optical microscope (20×). It can be observed that the printing layout circuit on the Cu foil surface facing to the cathode after the exposure was better than that printed on the surface facing to the anode and better than that printed on the surface before exposure. This confirms that the surface adhesion forces were increased after the exposure process.

4. Conclusions

A cold plasma experiment was designed, established and operated at Ain Shams University (Egypt) by specifications and characteristics within the normal range of previous DC glow discharge plasma experiments in different labs around the world. The plasma is generated under a low-pressure range (0.15 - 0.70 Torr) with a maximum dc power supply of 200 W. The experiment consists of three main parts—the discharge chamber, the vacuum system and the powering circuit.

The cylindrical discharge chamber is made of stainless steel and has two movable

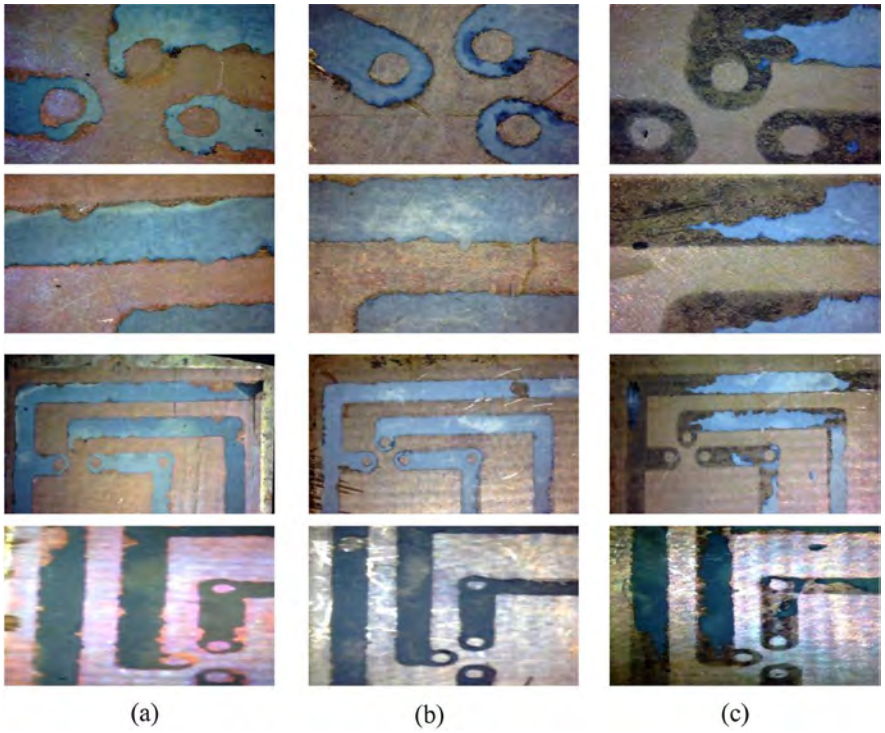


Figure 9. The circuit layout printed on the PCB. (a) Before exposure; (b) Facing to the cathode; (c) Facing to the anode.

cylindrical copper electrodes (cathode and anode), a vacuum system to evacuate the chamber before inserting the Ar gas at the operating pressure, an electric circuit that includes a power supply, potential divider, and high voltage probe.

To determine the operating conditions, the chamber was evacuated and filled it with Argon gas pressure 0.3 Torr. After that operating voltage V_{ps} was applied in the range from 400 volts to 1800 volts. The experimental work included measuring the electrical plasma parameters which described the electrical properties of the produced plasma adding to the electron temperature and ion density which were appointed by a double electric probe. From it, kT_e was in the range from 6.58 eV to 10.44 eV while n_i was in the range from 9.10×10^9 cm^{-3} to 1.79×10^{10} cm^{-3}.

After validating our experimental results, we used it as a cold plasma source and studied the effect of the generated plasma on the electrical and morphological characteristics of a Cu foil, such as the electrical conductivity and the surface adhesive forces using a printed circuit board (PCB). The aim of this work was to improve the electronic circuits' performance that can be controlled by improvement of the electrical characteristics of a PCB. The measurements show an improvement in the electrical conductivity of a Cu foil and the adhesion forces of the Cu foil surface after plasma exposure.

The resistance of the Cu foil (R) after being exposed to our lab generated cold plasma is found to be decreased by 61.1%, *i.e.* the electrical conductivity of the Cu foil is increased by 156.25% to allow greater current to flow. Thus, the di-

mensions (length/width) of the Cu trace that carries the electrical currents in the Cu foil in the electronic circuit of the PCB can be reduced. And in turn, the electrical conductivity (σ) of the Cu foil is increased. So, the electronic components can be increased with the smaller size of Cu foil on a PCB. The reduction in the Cu foil size decreases the cost of PCB manufacturing which is used in microelectronics applications.

On the other hand, the adhesion forces of the Cu foil surface after exposure were increased. This helps in decreasing the loss of the PCBs because of the defective printing the layout circuits on them. And this can contribute to the reduction of the price of electronic devices, which contain PCBs, to reduce the burden on simple societies.

Acknowledgements

I wish to thank and gratitude to staff in the central physics laboratory at physics department, Faculty of Science, Ain Shams University, for their great efforts in measuring the resistance of Cu foil in the PCB.

Conflicts of Interest

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

References

[1] Bellan, P.M. (2008) Fundamentals of Plasma Physics. Cambridge University Press, Cambridge.

[2] Bonizzoni, G. and Vassallo, E. (2002) *Vacuum*, **64**, 327-336. https://doi.org/10.1016/S0042-207X(01)00341-4

[3] Saad, M.M. (2013) Circuit Design of Glow Discharge Plasma Panel and Parameters Optimization. Thesis of M.Sc., Ain Shams University, Faculty of Science, Department of Physics, Cairo.

[4] Konuma, M. (1992) Film Deposition by Plasma Techniques. Springer, Berlin. https://doi.org/10.1007/978-3-642-84511-6

[5] Massarczyk, R., *et al.* (2016) Paschen's Law Studies in Cold Gases.

[6] Piel, A. (2010) Plasma Physics. Springer, Berlin. https://doi.org/10.1007/978-3-642-10491-6

[7] Kenneth Marcus, R. (1993) Glow Discharge Spectroscopies. Springer, Berlin.

[8] Podgornyi, I.M. (1971) Topics in Plasma Diagnostics. Springer, Berlin. https://doi.org/10.1007/978-1-4684-0724-2

[9] Uckan, T. (1987) *Review of Scientific Instruments*, **58**, 2260-2263. https://doi.org/10.1063/1.1139332

[10] Lieberman, M.A. and Lichtenberg, A.J. (2005) Principles of Plasma Discharges and Materials Processing. Wiley-Interscience, Hoboken. https://doi.org/10.1002/0471724254

[11] Masoud, M.M. (1989) Electric Probes and New Method. *Spring College on Plasma Physics*, Trieste, 15 May-9 June 1989, H4-SMR 393/69.

[12] Conde, L. (2011) An Introduction to Langmuir Probe Diagnostics of Plasmas. Madrid Dept. Física. ETSI Aeronáut ngenieros.

[13] Dahi, A., *et al.* (2017) *Separation and Purification Technology*, **187**, 127-136. https://doi.org/10.1016/j.seppur.2017.05.055

[14] Gao, S., Chen, S., Ji, Z., Tian, W. and Chen, J. (2017) *Advances in Mathematical Physics*, **2017**, Article ID: 9193149. https://doi.org/10.1155/2017/9193149

[15] Jafari, R., Asadollahi, S. and Farzaneh, M. (2013) *Plasma Chemistry and Plasma Processing*, **33**, 177-200. https://doi.org/10.1007/s11090-012-9413-9

[16] Castillo, H.A., Restrepo-Parra, E., De La Cruz, W., Bixbi, B. and Hernandez, A. (2017) *The Journal of Adhesion*, **94**, 615-626. https://doi.org/10.1080/00218464.2017.1320222

[17] Loureiro, J. and Amorim, J. (2016) Kinetics, and Spectroscopy of Low Temperature Plasmas. Springer International Publishing, Berlin. https://doi.org/10.1007/978-3-319-09253-9

[18] Jawaid, P., *et al.* (2016) *Journal of Cellular and Molecular Medicine*, **20**, 1737-1748. https://doi.org/10.1111/jcmm.12880

[19] Duquesne, S., Magniez, C. and Camino, G. (2007) Multifunctional Barriers for Flexible Structure. Vol. 97, Springer, Berlin.

[20] Edwards, P.R. (1991) Manufacturing Technology in the Electronics Industry. Springer, Berlin. https://doi.org/10.1007/978-94-011-3130-8

[21] Chen, C.-L., *et al.* (2016) New Surface Treatment Method for Copper Foils to Improve Signal Integrity of Printed Circuit Board. 2016 *Asia-Pacific International Symposium on Electromagnetic Compatibility*, Shenzhen, 17-21 May 2016, 584-586. https://doi.org/10.1109/APEMC.2016.7522804

[22] Suzuki, O., *et al.* (2016) Nano Anchoring Copper Foil for Next Generation Printed Wiring Boards. 2016 *Pan Pacific Microelectronics Symposium*, Big Island, 25-28 January 2016, 1-6. https://doi.org/10.1109/PanPacific.2016.7428430

[23] Weiss, B., *et al.* (2002) *Sensors and Actuators A*: *Physical*, **99**, 172-182. https://doi.org/10.1016/S0924-4247(01)00877-9

[24] Kotp, E.F. and Al-Ojeery, A.A. (2012) *Journal of Basic & Applied Sciences*, **6**, 817-825.

[25] Ghimire, B., Khanal, R. and Subedi, D.P. (2014) *Journal of Science*, *Engineering and Technology*, **10**, 20-27.

[26] Nojiri, K. (2015) Dry Etching Technology for Semiconductors. Springer International Publishing, Berlin. https://doi.org/10.1007/978-3-319-10295-5

[27] Hashmi, S., Batalha, G.F., Van Tyne, C.J. and Yilbas, B.S. (2014) Comprehensive Materials Processing. Elsevier, Amsterdam.

[28] Oks, E. and Brown, I. (2002) Emerging Applications of Vacuum-Arc-Produced Plasma, Ion and Electron Beams. Springer, Berlin.

[29] Trace Resistance Calculator—Electrical Engineering & Electronics Tools. https://www.allaboutcircuits.com/tools/trace-resistance-calculator

[30] Ganssle, J.G. (2008) Embedded Hardware. Elsevier, Amsterdam.

[31] Bolton, W. (2006) Engineering Science. Newnes, Oxford.

[32] Khare, P., Priyadarshini, M. and Verma, M. (2017) *International Journal of Engineering Technology Science and Research*, **4**, 30-36.

Journal of Modern Physics, 2020, 11, 1058-1074
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Creation Model from the Gell-Mann Standard Model to the Creation of Bio Cells: Based on the Assumption of Homogeneous 5D Space-Time Universe

Kai Wai Wong[1], Wan Ki Chow[2]

[1]Department of Physics and Astronomy, University of Kansas, Lawrence, USA
[2]Department of Building Services Engineering, The Hong Kong Polytechnic University, Hong Kong, China
Email: kww88ng@gmail.com

How to cite this paper: Wong, K.W. and Chow, W.K. (2020) A Creation Model from the Gell-Mann Standard Model to the Creation of Bio Cells: Based on the Assumption of Homogeneous 5D Space-Time Universe. *Journal of Modern Physics*, **11**, 1058-1074.
https://doi.org/10.4236/jmp.2020.117067

Received: June 15, 2020
Accepted: July 17, 2020
Published: July 20, 2020

Abstract

In this paper, we briefly go over the homogeneous 5D model field theory: from the 5D space-time inception, to its quantum field solutions given in terms of Higgs vacuum, filled with magnetic monopole bose fields of all energies. Then through the space dimension reduction projections, the Gell-Mann standard model was obtained as well as a quantum to Classical connection was made via introducing Bose distribution to the monopoles to obtain the Perelman entropy and Ricci Flow mappings. This provided us a picture to the creation of Astronomical objects, from galaxies to stars and planets. This method of splitting the monopole energy into ranges is extended to show that below the basic rest mass range of the electron and Quark, it still can be applied to explaining for the creation of the chemical elements periodic table. But perhaps the most interesting is in the lowest hundreds of Hz energy range, obtained from yet another 3 fold space symmetry breaking, into 2D × 1D, producing bio nitrogenous bases composed of 3 Carbon 12 in hexagon structures, due to preservation of the 1D monopole standing waves of this low frequencies. From that by imposing gauge changes the monopole states into DNA spectra. Since such spectra states retain the DLRO, it induces formation of charge carriers periodicity in a spherical bio cell.. It was then argued that due to cell's surface proteins, the structure must contain partial filled VB, with "p" state hole density, and empty CB, separated from VB by a positive band gap. Such band structures resemble known HTC Cuprate ceramics. Since the HTC goes through a Superconductivity transition via the simultaneous bose exciton condensation, providing a Coulomb pressure, which reduces the band gap substantially, and induces the ODLRO transition of the hole density. The same obviously applies to the bio cells. Because of the

near continuous exciton levels generated, a matching to the DNA spectra then can always occur by selective choices of proteins on the cell surface. Judging from a numerical study, we did years ago on YBCO, with doping. We found with a large enough VB hole density, the exciton induced superconducting gap can easily lead to Tc in the room temperature range. In fact by EMF excitation can increase the exciton pressure and trigger the ODLRO transition Tc upward. In fact, numerical results then suggest there do exist coherent EMF spectra from three key elements: Water, Carbon and Hydrogen, together with Oxygen, as studied over the years by numerous people, starting from Schrödinger to most recently Geesink.

Keywords

5D Fermat's Theorem, Space Dimension Reduction Projections, Perelman Mappings, The Higgs Vacuum: A B.E. Condensed Monopole Bosons, Realization of Excitonic Induced Superconducting "p" Valence Band Orbitals in Bio Cells, The Final Coherent Building Block EMF Spectra: Water, Carbon, Hydrogen and Oxygen

1. Introduction

Our advances in knowledge on mathematics, natural sciences, and biological sciences have accelerated since Newton's theory of gravity, based on principles that do not appear to be linked. In this paper, we hope to show that all these principles for all the areas of our knowledge are in fact linked through mathematical theorems associated with a homogeneous 5D common space-time topology. The complete mathematical understanding of the Electromagnetic Theory was forwarded by Maxwell, by recognizing that the 4 electromagnetic vector potentials are the solutions to that of the 4D homogeneous space-time, which obeys the Fermat's Last Theorem [1]: a seemingly very simple space-time topology. Yet the solutions between electric and magnetic fields are unsymmetrical [2]. This broken symmetry between electric and magnetic fields, led Maxwell to suggest a missing fifth component magnetic monopole potential, which can restore symmetry. Yet since, such a magnetic monopole was not observed within the framework of electrodynamics, the suggested magnetic monopole potential can then only form a direct product to those of the other 4 electromagnetic vector potentials. Hence, this monopole potential can only be the result of extending the space-time homogeneous manifold to 5D. Furthermore, since the 4D space satisfies Fermat's sum. It would necessarily make the radius "r" vector an entangled variable to the 4th space dimension. In fact, "r" as an entangled variable is in topological agreement to that of 4 space dimensions, as expressed in the Pythagorus sum, gives "r" with three angle variables: namely $0<\theta<\pi$; $0<\varphi<2\pi$ and $0<\psi<4\pi$. ψ limit happens to be identical to that of the solid angle for 3D space. It is based on this topological observation, and from the concept of having massless, e and −e spinor, we have derived the

explicit form of the Maxwell's magnetic monopole [3]. Together with the application of the uncertainty principle, and the space dimension reduction by projections into a 4D Lorentz space-time, we have then extended the homogeneous 5D as the space-time of the universe, with the reciprocal 5D energy, momenta representation. From which we presented a grand unified field theory that encompasses the Gell-Mann quark model [4], as obtained from the space to space conformal projection P1, and the Leptonic weak theory, an equivalent to separating e and −e of the charge neutral magnetic monopole through the space to time projection Po into only gauge conserved −e massive spinor [5]; while the space to space conformal projection that gives the fractionally charged Quark model that, for the +e component, must then retain and satisfy gauge invariance. Hence, the Higgs' vacuum [6] as identified in terms of the DLRO (Diagonal Long Range Order) Bose-Einstein condensed monopoles massless, charge 0, Bose fields with arbitrary quantum energies $2h\nu$ was developed [3].

It is the purpose of this paper to both summarize and fill in gaps left not addressed in reference [5]. Thus readers that did not read [5] can find the rest of the paper incomprehensible.

As the 5D homogeneous manifold is chargeless, massless, and does not possess time reversal symmetry, all space dimension projections must retain charge neutrality at all time in the reduced 4D Lorentz domain. Therefore, when an electron is created, a proton must also be created. This charge neutrality constrain in the presence of a Coulomb potential, and via the Schrödinger's equation, leads us to the Bohr model, for atoms, and gives us the periodic table for elements, provided the protons are in the respective nucleus. Since the proton is obtained from the combination of the up, up and down quarks confined by gauge, the protons within a nucleus must also be gauge confined together. That means the gauge loop within the nucleus must be given by a N order knot, looping over all N nucleons within the nucleus. In 3D space, a closed loop uniquely defines a plane in which it encloses, thus a knot can be transformed into a single loop.

Under 4D space, such a knot cannot be transformed into a single 2D loop, as there remains two variables, such that a single 2D plane is not uniquely definable. It is this mathematical topology that allows the gauge invariance to holds the repulsive protons within the nucleus together, similar to the gluons which are tensor products of the EM vector potentials under gauge confinement give raise to most of the nucleon rest mass. Based on the above, the homogeneous 5D quantum field Theory was forwarded by us to explain the origin for the universe, including proving the quantum to classical Perelman-Ricci-Flow and Perelman-entropy mappings [7] [8] connections. Through this connection provides us with a picture on how galaxies, stars and planets were created during the Big Bang [9] [10]. We will further discuss in section 2 more details on the mathematical relations between the quantum and classical mappings from 5D to Lorentz 4D.

There remains then, the very low monopole energy region, that is

$0 < 2h\nu < 2m(e)c^2$ region. It was shown previously that the monopole boson field can exist within a carbon 12 nuclear core, provided it can complete a closed gauge loop via quantum tunneling from carbon core to carbon core. To achieve this gauge loop we had proposed that the process of quantum tunneling is from a nitrogenous base to the next layer nitrogenous base with random ordering and reflected back at the end caps base [11], hence producing the DNA spectra of the life form. This very last low frequencies monopole being of DLRO spectra of the DNA is than applied as a global ordering induction, so that the ODLRO (Off-Diagonal-Long-Range-Order) phase transition of the charge carriers within the bio cells occurs, thereby leading to a natural cell growth [12]. It is this low frequency quantum feature of the monopole states that produces the EMF coherent and de-coherent quantum ensemble properties that were first suggested by Schrödinger as the bases of life and extended quantitatively by Frohlich [13]. In the forth section of this paper, we follow the EMF expression given in the recent works of Hans J. Geesink and his collaborators, and examine how their expression on the EMF spectra of biological systems [14] [15] can be made compatible to that from the 5D field theory.

2. Big Bang

Most astronomers believe in the Big Bang for the beginning of the Universe. It is fascinating to investigate the meaning of this Beginning of Creation.

We had postulated that the Universe is in a homogeneous 5D space-time manifold [6], which satisfies the Fermat's Last Theorem [1]:

$$(ct)^2 = \sum_{j=1}^{4} x_j^2 = r^2 \tag{1}$$

r is the Fermat's amplitude, and r^2 is a positive scaler, spanning "an increasing" space volume.

In fact, this homogeneous 5D manifold has several important physical meanings. Namely: As space is only positive, time can only increase. As the negative "t" root from the quadratic Equation (1) violates causality, when r^2 spans "the increasing" volume. Hence t only increases. This remark is also valid for the homogeneous 4D Maxwell manifold, giving us that light only propagates outward in time from its source. Thus any fields within the manifold do not have time reversal symmetry. Secondly, the homogeneous manifold contains no net amount of charges of either sign, which means all creations from this homogeneous space-time must maintain net charge neutrality. Thirdly, in order that matter can be created from it by space dimension reduction into a 4D Lorentz manifold, where Einstein's $E = mc^2$ is satisfied, there must exist a corresponding Fermat's representation in terms of Energy and 4 momenta:

$$E^2 = \sum_{j=1}^{4} c^2 p_j^2 \tag{2}$$

Because of Equation (1) and Equation (2), there must exist a unique relationship between time variation to Energy variation, which of course would also de-

fine the same variation relationship between space and momentum. This relationship is the uncertainty principle:

$$\delta E \delta t \geq h \tag{3}$$

Since the Universe started from $t = 0$, implies it must start from infinite amount of energy. Therefore, it started with a Big Bang.

Now applying the uncertainty principle, the Fermat's sum of Energy and momenta, becomes a second order homogeneous quadratic differential operator, which means solutions within it must be massless, and 0 net charge waves, but they cannot be electromagnetic waves. As electromagnetic waves are in the homogeneous 4D Maxwell space-time, their existence require the present of an electrical current source.

Hence, the massless wave within the 5D that carries energy only, must be a Boson field with net 0 charge, and according to Maxwell [2] generates the magnetic monopole potential, namely:

$$+ \text{ or } - \frac{2ec\mu}{r} \tag{4}$$

where μ is the magnetic susceptibility.

Where $+e$ and $-e$ are the charge signature of massless spinor, that are solutions to the Dirac linearized differential equation obtained from Equation (2). These massless charged spinors were coined by us [5] as e-trino, and anti-e-trino. Therefore, Equation (4) represents the radiating current at speed c along the Fermat's radius "r", by a net charge 0, and 0 momentum Boson field obtained from the product of a pair of opposite momentum e-trino and anti-e-trino, which carries a quantum energy of $2h\nu$. The restriction that this opposite charged massless spinor pair must also have opposite momenta, is because they must be in a Diagonal Long Range Order, and filled the expending homogeneous 5D manifold, so that in energy, momentum representation it preserves homogeneity also. It is this very important feature, that we can consider that the homogeneous 5D Universe, must be given by a Bose-Einstein condensed vacuum with energies carried by e-trino, anti-e-trino pairs. In terms of Higgs' theory [6], these bosons are the Higgs' fields.

It is also interesting to observe that $2ec$ looks like a charge current, and if so should generate electromagnetic waves as represented by photons, yet photons radiation are given as Poynting vectors, which has orthogonal E, H oscillating fields. While as monopole solution, E, H cannot be present. Therefore, in terms of photons, this Boson if it radiates photons, must radiates a 180 degree off-phase photon pair, having energy $2h\nu$ and momentum $2h\nu/c$.

However, this Higgs' Boson is in the BE condensed state, it cannot have a net momentum. Hence, the $2ec$ is not a charge current, and c is just a parameter not the velocity of photon, hence $2ec$ represents a magnetic monopole value and does not radiates [3].

$$J(r) = ec - ec = 0 \tag{5}$$

This magnetic monopole potential given by Equation (5) is not of the same dimension as the Coulomb potential. While the magnetic monopole potential energy is in the same dimension as the Coulomb potential energy.

In electrodynamics, given in the homogeneous 4D Maxwell manifold, energy is carried by photons emitted from a charged current, which as we mentioned is absent in the homogeneous 5D manifold, hence although the monopole Bosons carry energy, we some time describe this 5D space-time of the Universe as Dark Matter domain, a vacuum filled with the Bose-Einstein condensed Higgs' bosons. Accordingly, we can define a Bose distribution

$$\frac{1}{\mathrm{e}^{\frac{h\nu''}{kT}}-1} \tag{6}$$

where $\nu''=\nu'-2\nu>0$, and is the frequency of the excited monopole with eigen-energy $h\nu'$.

Therefore the vacuum is excited when $h\nu''>0$, and the BE condensation is broken. We can likewise view the vacuum as providing a quantum potential well:

$$W=-4\pi(2ec)^2\mu\frac{(kTN)^2}{cr} \tag{7}$$

where

$$N=2\sum_{j=1}^{\infty}\frac{\mathrm{e}^{-jC}}{j} \tag{8}$$

and $C=2m(e)c^2/kT$.

N comes from integrating $h\nu''/kT$ from C to infinity. Note N is finite as long as $C>0$. This quantum well then represents the energy converted into massive $-e$ charged lepton and +e massive proton. ν'' being a frequency above 2ν of a specific Higgs' field is critical to holding the creation of a massive electron to that required to simultaneously create a proton and possibly also neutrons by the separation of e and $-e$ in the monopole, such that charge neutrality can be maintained at all time t, despite the proton and neutron creation needs the creation of the SU(3) generators: $2/3e$; $2/3e$ and $-1/3e$. Hence to be simultaneous must be obtained from at least different Higgs' field [see chapter 8, of reference 5 for detail]. It should be emphasized that the creating of a complete set of quarks requires energy equivalent to $m(q)c^2$, where $m(q)$ is the bare quark mass.

Since the quarks are obtained by the conformal projection P1, from the 4th space dimension onto any other remaining 3 space dimensions. With 3 quark generators, there are therefore $4\times4\times4$ different choices of the P1 projections. Or that we require at least 64 different energy Higgs' bosons. Each contributes at least $m(e)c^2$ due to the 3 P1 choices to create simultaneously the proton. Hence the ratio of $m(q)/m(e)=4\times4\times4$. In another word $m(q)c^2=64\times m(e)c^2$, which is 32 MeV. In agreement to the value we obtained from the experimentally measured and model calculated hadron masses [5] [16] [17]. It is this rela-

tionship that makes the minimum C value depends on twice the electron rest mass energy to be able to maintain charge neutrality at all time.

Therefore within the 4D Lorentz domain, the total ensemble averaged energy of the projections created masses satisfies

$$\langle E \rangle_T + W = 0 \tag{9}$$

This new law allows the 5D Universe to start with infinite energy as is given by the Big Bang, and simultaneous provide with a negative infinite W, when no mass is present C was 0, when $t = 0$, giving us that the Universe creation is from a very philosophical absolute NOTHING, with no space and no time.

Based on Equation (8) we have published an article [11] connecting this thermal quantum mass creation to those of Perelman's Ricci-Flow and Perelman's entropy mappings [7] [8] and applied it to the study of stars and planets with good successes [9] [10].

Giving us a further topological understanding on the processes and time sequence of different sizes galaxies, stars and planets creation in their respective 4D Lorentz space-time domains out from the 5D homogeneous space-time manifold.

As masses moves within the Lorentz space-time, their energies and momenta must be expressed in covariant form, hence leading us to formulate the Lorentz domain in terms of covariant Riemannian curvatures, thus establishes Einstein's General Relativity [18]. It is from the curvature metrics that we derive the non-linear gravity equation and deduced the force of gravity between masses. As this gravity equation is non-linear, its solutions contain singularities in the Lorentz manifold. It is such singularities that become sinks for sucking in matters, unless either artificially excluded from containing masses, or else a counter repulsive force is automatically generated when masses spiral into these singularities [19]. Without which, masses will have to revert back into pure energy. In the 5D theory [5], all masses originated by irreversible space dimension reduction projections. As the homogenous 5D manifold has no time reversal symmetry, that means mass cannot revert back to pure energy. Thus the suggested automatic removal of the gravitation singularities must exist. In fact, from the Perelman's entropy mapping, we observed that for massive systems having more than 12 nucleons, there must exist a time frozen 4D space void core. This topology gives us the lowest spherical nucleus, the carbon 12 nucleus. Such a C12 nucleus, would have contained monopoles with standing waves energies higher than $2m(e)c^2$, thus not possible as we discussed in ref. [11]. But then nucleus do not collapse due to proton-proton repulsion. While astronomical objects, like stars, and planets, the void core is large, and a self-rotation is induced by transforming the radial monopole $2ec$, to that of e, and $-e$ moving in phase thus producing an angular momentum which in turn must be canceled by the self-rotation of the mass shell, as the original 5D domain contains no momentum. It is this self-rotation that will provide a natural centrifugal force to prevent the masses to collapse into the gravity singularity [9]. But none-the-less, a very

massive star will still experience the gravitational inward force, thus resulting in a supernova. When that happens, the transforming of the monopole states within the core into angular momentum states, must break the DLRO of the monopole. This breaking changes it to two in phase with opposite charge sign circular currents. Such a pair of circular currents, then generate radial radiating pairs of photons with 180 degree off-phase, resulting in a $2h\nu/c$ graviton. In fact, this supernova graviton radiation from Supernova was observed [19] [20].

We have discussed the monopole with energy above $2m(e)c^2$, which are convertible into electrons and nucleons, yet conserving charge neutrality in the 4D Lorentz domain. That leaves us the low energy range $0 < 2h\nu < 2m(e)c^2$. In the analysis of the excitation from the Higgs' vacuum filled with Higgs' DLRO Bosons with all energy values, that the simple Bose distribution from the excitation of the vacuum will result in N being infinite at the lower limit $h\nu'' = 0$. Physically, this means the vacuum cannot be excited, at any temperature T. By that it can only means all such 1D monopole states must be time frozen, and is a standing wave, or that it is in a closed loop, being achieved by two parallel and opposite wave propagation. It is the first one, that allows for such a standing wave to exist within a time frozen 4D space void core. For a carbon 12 nucleus, its dimension restricts the monopole frequency to exceed the electron rest mass correspondence. Hence, in order to accommodate the range we need quantum tunneling from C to C along a line to happen, which is okay, as the monopole is a DLRO Boson. Such a process for accommodating low frequency monopole states can be achieved by introducing a set of 4 nitrogenous bases. The fact that there is 4 and only 4 such possible bases, came from in obtaining a 2D plane out of a 4D space, we have 4!/2! = 12, which must in turn obey the remainder 3D space orientation invariance. The length of the nitrogenous bases stacking out of all possible orderings, then gives us the structure of a RNA, with one of the nitrogenous base identified as the end cap, which two faces have different properties: namely, transmitting and the other non-transmitting. Since the stacking can be separated arbitrarily with none carbon base layers in between, such a DLRO monopole wave can have frequencies within the low range we desire. But it however does not produce a quantum flux of $h/2e$. On the other hand, if we combine two parallel RNAs, but replace the end cap nitrogenous base to one that not only reflects on one side but also reflects the monopole wave into the adjacent RNAi. The index "I" is meant to signal the end cap replaced. Under this result, we form a DNA. Although the choices of stacking are fixed for all DNAs, but due to the infinite choices of the separating gap layers, there are infinite different spectra possible. And this monopole wave completes a closed loop, hence always gives us the quantum flux $h/2e$. The flux quantization is vital, if cells in life forms derive its growth from matching the DNAs spectrum with the charge carriers in the cell go through an ODLRO superconducting phase transition [21].

Since the free charge carriers in a solid requires a periodic elemental structure, so that conduction and valence bands can be computed. One can generally

group the carrier properties as follows: 1) metal, where there is a partially electron orbital filled CB conduction band. In short, the Fermi level cuts through the conduction band. 2) A hole carrier conductor. There the Fermi level cuts through the VB the valence band, leaving a net fraction of VB orbitals vacant. The Cuprate high Tc ceramics belongs to this class of conductors. 3) A completely filled VB and a completely empty CB, separated by a small band gap energy G. Such a system is a semi-conductor, when electrons from VB are excited into CB. 4) When G is large compared with excitation energy available, it is an insulator. Actually even when excitation can be activated, the opposite charges left behind in VB and CB can bind by Coulomb potential and thus form a discrete exciton level within the band gap G. It will still be an insulator. 5) For the Cuprate ceramic having a partially filled VB [22], but still can excite a VB electron into the empty CB, and formed a localized exciton level within the band G. Such an additional localized atomic like exciton state adds to those of the Crystal ions, and hence increases the induced coupling between the hole carriers to form ODLRO Cooper pairs [23], thus substantially enhances the superconducting transition temperature Tc. Bio cells grow at a room temperature range. And such bio cells are usually consisted of a spherical shell, predominantly made of different proteins. Because of the spherical shell structure, bio cells all satisfy the periodic boundary condition, thus a band structure having distinct CB and VB must occur. Since all proteins contain carbons arranged in hexagon form. Such a hexagon coupled with carbon having outer atomic orbitals $n = 2$; 2p and 2s, the 3 nearest C neighbors in fact will leave a partially filled VB "p" orbital band. A result is quite similar to the Cuprate HTS ceramic. The differences could very well be just quantitative. In YBCO, we have observed experimentally, that Tc can be substantially enhanced with UV light irradiation [24], which implies that the exciton within G, can be excited so that its binding raises closer to either CB or VB, depending on the relative effective masses of the CB and VB excited extra charges. It is this effect that makes the exciton polarization easier, we may then expect cells to drive its Tc transition into the room temperature range when EMF (Electromagnetic wave frequencies) are make available, internally in animals via neural spectrum, or in plants through photo synthesis. It is such EMF induced Tc phase transition that we consider as the bio coherent frequencies. EMF that do not induced the Tc transition in the cell structures are than classified as de-coherent. Many materials including common clay do process a range of EMF overlapping with both coherent and de-coherent frequencies. And it is not unexpected, as plants derive and form its proteins from absorbing from the soil which is of course clay dissolved in water.

3. The Chemical Periodic Table

The chemical periodic table is based on the Bohr's model. To clearly understand let us consider the simplest two atoms H and Li. The atomic orbitals for H is 1s, and for Li is 2s. For Li, we have 3 protons and 2 neutrons in the nucleus. There-

fore, we must have 3 electrons also. Note all stable atoms are charge neutral and are Bosons, because they must be composed of even number of Fermions. If we ionized Li by removing the 2s orbital, the change in the atomic binding is $-R/2^2$, corresponding to an energy gain of $[1/4]R$. While if we consider the binding of the nucleus to the remaining 2, 1s orbitals after the outer 2s is removed, we have $2 \times [-2R] = -4R$. The factor $-2R$ comes from the effective nucleus charge seen by each 1s electron, after the outer 2s electron is removed, thus changing the Coulomb potential between each remaining electron with the nucleus. As compared to the non-ionized case, where we simply have $-2R$, due to Pauli Exclusion principle for the 2 1s electrons, since we must treat all 3 electrons as seeing the same Coulomb field, otherwise they are distinguishable, and the outermost orbital would not be 2s. There is an extra binding of $-2R$. This extra energy makes the total energy gain of $[9/4]R$. Hence resulting in a net change of $[9/4]R > [1/4]R$. Furthermore, the ionized Li is no longer a Boson, unless the nucleus lost a neutron also. The initial homogeneous 5D quantum representation only allows for a DLRO Higgs vacuum, making the retaining of Boson fields under excitation a quantum condition. According to thermal ensemble equilibrium condition: that always minimizes the energy value, the ionized Li state is unfavored. Including the Boson requirement, the elements in the periodic table must be formed right below the leptons, hadrons creation. With the energy shift dominated by the number of nucleons in the element. Meaning the atomic formation favors charge neutrality, and which means atoms are formed under simultaneous charge neutrality as electrons and nucleons are formed by the projections Po and P1, from the homogeneous 5D manifold into the 4D Lorentz manifold where they exist. It is interesting to see that this thermal energy difference corresponding to a value of $2R$ is an unfavored photon EMF shift of extra $2R$, as compared to the favored EMF shift of $[1/4]R$. Therefore this interesting charge neutrality condition can be seen as dividing the EMF spectra into coherent and de-coherent values in the chemical elements formation. Except for the chemical elements, the EMF spectra is more than two order higher than those of the bio cells molecules, such as water, given by Hans Geesink's papers. In fact the periodic table elements creation is distinctly a separate energy range below the leptons and hadrons formation obtained via the space reduction projections Po and P1, yet far above the bio EMF range. Making the presents of life forms on planets very restrictive and rare. Nonetheless the same coherent, de-coherent analysis for the much lower EMF spectra still remains, and applicable to the creating of bio-molecules, when the projected 4D space to 3D space symmetry can be further reduced to 2D × 1D space as is the space symmetry for RNA and DNA. This restricts the planet's surface Temperature to a very narrow range governed by the liquid phase of water, as water together with carbon 12, and Hydrogen, in replacement to proton, neutron and electron in atoms, due to fractal representation play the most important key roles in the maintenance of Bose monopole DLRO for the bio EMF coherent spectra. {We like to mention the mass energy for Hydrogen as compared to water is 1/18, while the electron to

the proton is $1/64 = 1/4^3$, coming from 4 choices to select the 4th space coordinate to be projected onto the remaining 3 coordinates, and the power 3 is from the symmetry of the 3 remaining space coordinates. The increases of $(2/3)^2 \times 2^3$ times as the electron is replaced by Hydrogen, is due to breaking of 3D space symmetry to 2D × 1D, starting from a 5D homogeneous manifold. The 2/3 factor comes from a 2D plane selected from 3D space; the square factor is from the quadratic Fermat's sum, while the 2^3 comes from the three orthogonal normals to the planes. We still need to define the 2D geometric form. On the carbon 12 nucleus surface, the form is given by the 3 hexagons. Each hexagon is formed by 3 equilateral triangles on one side as given in the nitrogenous bases, which in terms of the SU(3) generators or carbon is given by $(2/3)^3$, so that the DLRO of the monopole bosons can be linked via quantum tunneling in order to extend the mapping to lower EMF values; while we have 3 such equilateral triangles, hence we have exactly $(2/3)^3 \times 3 = 8/9$; which then gives us Oxygen. As a result all bio lives are built from water, carbon, hydrogen plus oxygen.} Such Low EMF spectra derived for bio systems was first suggested by Schrödinger, and now extensively analyzed by many bio-physicists. The EMF obtained from water is based on the molecular rotation of 432 Hz, which is available only in the liquid phase, thus lives are limited to the water liquid phase Temperature range. In agreement with our Bose distribution approach as we divide the EMF value into separate ranges so as to connect the projected space symmetry to that of Perelman's Ricci-Flow and entropy mappings, and limiting the thermal requirement on the existence of Life forms on a planet. It is this stringent limitation that made Earth a rather unique planet within the solar system.

4. Analysis of the Bio-Cell EMF Spectra According to the EEM Theory for Its Growth

In a recent article by Geesink, Jerman and Meijer [25], the bio-cell's coherent EMF is given by

$$E(n,m,p) = hw(ref)2^n 3^m \left(2^p\right) \tag{10}$$

where $w(ref)$ is a cell specific frequency.

$$n = 0, 0.5, 2, 4, 5, 7, 8, -1, -3, -4, -6, -7.$$

$$m = 0, 1, 2, 3, 4, 5, -1, -2, -3, -4, -5.$$

$$p = <-4, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, > 52.$$

We have proposed a 5D model theory for Life [12] based on the EEM model [22] for inducing an ODLRO transition of the VB "p" holes within the bio cell, as it matches to the DNA spectra, which obey gauge quantization $h/2e$ from the DLRO monopole bosons in the life form. In this section, we shall try and explain why the EMF spectra given in Equation (10) are consistent with our model theory. First it is nice to observe that the EMF corresponding to the masses of H_2O and Carbon ratio is exactly 1 to 2/3; (EMF of water: 432 Hz, see [25] for the EMF values) the same ratio as between proton and neutron quark masses, while

hydrogen fits exactly according to the electron to proton ratio computation, which is exactly $m(\mathrm{H}_2)=m(q)\times 1/4\times 2/3\times 2/3$. The factors 2/3 × 2/3 comes from the SU(3) factor for carbon, and the remaining 1/4 comes from the choice of 4D space in Po projections, and reduced the 3D space symmetry into 2D × 1D as required by the formation of DNA. $m(q)$ corresponds to the EMF frequency of Water. As such, these 3 elements, water, carbon and hydrogen are the only basics that are in all bio matters; as only proton, neutron and electron are for the entire periodic table of chemical elements. Oxygen however, would correspond to 8/9 of water EMF. This is not a basic SU(3) derived fundamental. However it does correspond to 3 carbons in product geometry, a very interesting arrangement present in all the nitrogenous bases. Perhaps that makes Oxygen vital to the creating of spectra matching between those from the DNA and the ODLRO of carriers in the bio cells?

In order that an ODLRO transition on the bio cell's hole carriers can occur in match to the DLRO monopole bosons spectra for the gauge loops of the DNA, a band structure with empty CB and partially filled VB with "p" holes, must be established for the bio cell. All bio cells have a spherical surface covered with proteins. Thus the surface satisfies the periodic boundary requirement for the existence of a band structure. As all proteins contains hexagonal carbon structures, and with the n = 2, 2p, 2s outer orbitals for Carbon, the hexagon structure will give rise to a partially filled "p" orbitals as the Fermi level cuts through the VB, and depending on the rest an empty CB, separated from VB with a positive band gap G. Such a band structure resembles the YBCO HTC band structure [22]. The physical meaning from such a band structure, is that in the normal phase, it is a hole type conductor. Except this band structure might not be the lowest energy configuration. Because when an extra VB electron is excited into the empty CB, it requires $Ef+G$ energy, while these pair of opposite charge sign particles will bind with Coulomb potential forming a bound exciton state. Such an exciton brings down the excitation level to $Ef+G-E'(ex)$. Depending on the relative masses of the CB electron and the VB hole. And as $E'(ex)$ cancels G, the exciton will disintegrate. Thus there are two possibilities, one from $E'(ex)$ to be just above VB, and two for $E'(ex)$ to become 0, or at the bottom of CB. We may denote these two conditions by an absolute energy $|dE'(ex)|$ measured from either the VB top, or the CB bottom. Obviously $|dE'(ex)|$ is given by $hw(ref).F(n)$, where $F(n)$ is a function of the quantum orbital levels of the exciton. On the cell's surface, the exciton can neither be like a 2D Bohr atom, or a 3D Bohr atom. Therefore more accurately $F(n)$ is a combination of $F(n;m-1/2)$, and n, m are integers excluding 0. Such that

$$E'(ex)=a\frac{R}{n^2}+b\frac{R^*}{(m-1/2)^2} \tag{11}$$

where n, m are ± integers, except 0. R and R^* are the respective Rydberg constants. a, b are the dimension mixing. The above $E'(ex)$, can be rewritten as

$$E'(ex)=4\left\{\frac{aR}{(2n)^2}+\frac{bR^*}{(2m-1)^2}\right\} \tag{12}$$

The correct $E'(ex)$ in the EEM theory is given in Equation (6) of ref. [26].

This expression nonetheless shows that the $E'(ex)$ levels are a mixture of even and odd integers. However they are NOT the n, m powers shown in Equation (10). We use the same notations just to show that there is always a match to the DNA spectrum possible. They simply represent two infinite sets of DLRO bosons that can be coupled to the "p" hole carriers to induce the ODLRO transition, so as to match the DNA spectra. We refer to ref. [26] for a more accurate discussion on $E'(ex)$. As excitons formed, an internal Coulomb pressure arises, which collapses the band gap G, from eV to only order of 10^{-7} eV. The smaller the effective G^* is, the larger will be the induced ODLRO $E(ex)$ becomes. The photon frequency corresponding to $G^* = 10^{-7}$ eV. is roughly 10^{-3} G^* according to the numerical fitting results in [26], hence in the 100s Hz range, matching the data shown in [25]. In fact, as G^* is so small, the discrete $E'(ex)$ levels becomes continuous and these excitons mathematically becomes similar to the monopoles within the Higgs' vacuum, with frequencies in the 100s of Hz. And the induced hole-hole pairing obeys the SU(3), SU(2) symmetries exactly like the elements periodic tables we obtained from the Gell-Mann standard model. And the function $F(n;m-1/2)$ becomes $F'(2n;2m+1)$, where n, m are all plus and minus integers. Since F represents $|dE'(ex)|$ it represents the EMF induced infinite possible number of $E'(ex)$ level to level transitions. Hence $F(n, m)$ with matching to the DNA can become roughly $Eo2^{-n}3^{-m}$. As any integer powers of 3 remains odd, and all but 0 powers of 2 is even, thus satisfies $E'(ex)$.

The EEM induced ODLRO binding energy gap $E(ex)$ which is due to the energy saving by the hole-hole pairing is thus given by matching those of the DNA:

$$E(ex)=\frac{Ef^2}{F(n,m)}\cong hw(ref)2^n3^m \tag{13}$$

See [27] for $E(ex)$ solution.

The $Ef.Ef$ term comes from the energies of the hole pairs on the Fermi surface, while $F(n, m)$ is in the denominator is a second order perturbation correction to the total Cooper pair energy. Due to 5D projection, $E(ex)$ match to the DNA spectra must retain the SU(3) and SU(2) geometry symmetry, as we saw in Water and Carbon. Of course, $E(ex)$ is only part of the Cooper pair binding gap. The total binding gap must include the phonon induced BCS term. Giving us the total superconducting transition gap

$$E(S)=\left\{E(\text{phonon})^2+E(ex)^2\right\}^{1/2} \tag{14}$$

The phonon induced gap E(phonon) does not have quantum levels and is a pure constant, which we can assign as the power $n = 0$ from the 2^n EMF spectra. With these two orthogonal sum terms for $E(S)$, and the exciton state is domi-

nated by the 2D atomic spectra, with the 3D fraction $a \ll b$ and also depends on the level n, plus the + and − series difference due to the CB and VB carrier masses in assigning $F(n; m)$ we can get to the same results as observed by Geesink *et al.* except for $n = 0.5$. Such a spectrum can only be associated to the E(phonon) gap component as applied to the CB electrons, as the 3D exciton ground state energy behaves like the ionic mass in the BCS electron superconductor. The Tc depends roughly as a function of Nf^*/Nf, where Nf is the Fermi level hole density before the excitation of creating exciton and Cooper pair. The smaller this ratio, the higher the Tc. It is experimentally verified as the Tc isotope dependence [26].

The highest Tc found in HTC ceramics is around 125K. In order we have Tc in the room temperature range requires a 50% increase from the ceramics. Hence this can occur if the water rotation states given in Equation (10), in T above freezing contributes at least equally to that from the Excitonic gap. Implying water EMF plays a significant role in inducing the growth of bio cells.

Lastly, the EMF series (2^p), with symmetry between + and − p integers could very well be from the rotational states of water molecules on the cell, that can separately induced a Cooper pairing of the VB hole carriers, without the formation of a localized exciton. And due to second order perturbation correction we get a third E(S) gap enhancement, on top of those of the lattice phonon coupling, hence is nearly 3D symmetric.

A completely microscopic superconductivity field theory based on EEM was published and compared extensively to results of many different HTC cuprate ceramics [22] [23] [26]. It was also found that by UV light irradiation on a YBCO thin film, and creases Nf, and the Tc value is substantially increased [24]. This result and the EMF (2^p) series shows both the Tc enhancement to bio cells, leading to growth by the presence of EMF coming from photon energy within the life form body, and the presence of water molecules on the bio cell are essential.

5. Conclusion

We have briefly reviewed all physical results derived from the dimension reduction mappings from the homogeneous 5D space-time Universe into 4D Lorentz space-time domains where all matters exist. Starting with the creating of the BE condensed Higgs' vacuum, filled with Maxwell's Bose magnetic monopole states, to the realization of the electro-weak leptons, the Gell-Mann Quark model with gauge confinement has led to the 8-fold way hadron results, and proved the McGlinn theorem [28] to be valid and compactable when matters were derived from the 5D projections onto the 4D Lorentz space-time domain. And then by applying the thermal ensemble mathematically, we bridge the quantum picture on mass creation into the classical Perelman-Ricci-flow and entropy mappings, so as to obtain the creation of galaxies, stars and planets within each separated 4D Lorentz domain. The point we missed in our previous papers on these topics is the mathematical importance that monopole states are of DLRO and are along

"r" of the Fermat's radius, an entangled dimension. Therefore, in the homogeneous 5D, we have a vacuum filled with BE condensed monopole states, as suggested by Higgs. Because of this entangle variable, it is possible to retain charge neutrality at all time, when projections, or excitation of the monopole fields were applied. As negative electrons are simultaneously created with equal number of positive protons, it fixes the bare quark mass at 32 MeV. As compared to the electron rest mass of 0.5 MeV: a result not deducible from the Quark model or is expected. Through these results, we learn that all creations including lives are derived from excitation of the 5D Higgs' vacuum of BE condensed DLRO monopoles, into DLRO matter structures composing of electrons and nucleons, making up elements of the periodic table. While for even lower energies, elements condensed into solid structures, which converts its atomic orbitals into band structures, and nucleus motion into vibrating phonons. Eventually at even lower monopole energies, the excitations form 2D hexagon carbon structures, including clay, fats, proteins, and leading to nitrogenous bases. Through stacking these bases in all possible ways produces RNA, allowing the DLRO monopole field via quantum tunneling to satisfy the boundary condition of the low frequencies states. Similar to the creating of nucleons from quarks by imposing gauge invariance, the DNA for life forms emerges. Because of all these natural mapping steps, the 5D homogeneous energy-momentum manifold is divided into different ranges as a function of temperature T by keeping the $2h\nu(o)/kT = C$, a set of constants in satisfying the quantum to Perelman mappings relation. It is this $\nu(o)$ spectra and its gauge quantized restricted condition, that separates into coherent and de-coherent for life forms, that is being studied by Geesink and many others. The coherent spectra are the spectra responsible for cell growth. And based on the EEM model we suggested, it must then be governed by the bio charges, which are mainly unfilled "p" orbitals derived from carbon in the hexagon structures of proteins covering all cells, thus related to the photon energy induced exciton spectra, in the band gap that creates the ODLRO phase transition of these "p" hole carriers, matching those of the DNA, hence giving us the cell growth, as we discussed. The part that we have not discussed is how the photon inducing source comes about in the bio systems. For vegetations, it is photosynthesis. For animal lives, it has to be from internal. Obviously, there are much more possible channels. Error in the spectra generated will result in cancerous growth. Again its causes are many, from the DNA to the neuro-signals.

Acknowledgements

We like to thank Professor Hans Geesink for many private communications and providing us with his recent preprints. And we sincerely thank Ms. Winnie So for typing and careful review of the manuscript.

Conflicts of Interest

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

References

[1] Aczel, A.D. (1997) Fermat's Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem. Penguin Press, London.

[2] Maxwell, J.C. (1865) *Philosophical Transactions of Royal Society of London*, **155**, 459-512. https://doi.org/10.1098/rstl.1865.0008

[3] Wong, K.W., Dreschhoff, G., Jungner, H., Fung, P.C.W. and Chow, W.K. (2018) *Physics Essays*, **31**, 493-495. https://doi.org/10.4006/0836-1398-31.4.493

[4] Gell-Mann, M. (1964) *Physical Review Letters*, **12**, 155-156. https://doi.org/10.1103/PhysRevLett.12.155

[5] Wong, K.W., Dreschhoff, G. and Jungner, H. (2014) The Five Dimension Universe: A Creation and Grand Unified Field Theory Model. Scientific Research Publication.

[6] Higgs, P.W. (1964) *Physical Review Letters*, **13**, 508-509. https://doi.org/10.1103/PhysRevLett.13.508

[7] Perelman, G. (2002) The Entropy Formula for the Ricci Flow and Its Geometric Application.

[8] Perelman, G. (2003) Ricci Flow with Surgery on Three-Manifolds.

[9] Fung, P.C.W. and Wong, K.W. (2015) *Journal of Modern Physics*, **6**, 2303-2341. https://doi.org/10.4236/jmp.2015.615235

[10] Fung, P.C.W. and Wong, K.W. (2017) *Journal of Modern Physics*, **8**, 668-746. https://doi.org/10.4236/jmp.2017.84045

[11] Wong, K.W., Fung, P.C.W. and Chow, W.K. (2019) *Journal of Modern Physics*, **10**, 557-575. https://doi.org/10.4236/jmp.2019.105039

[12] Wong, K.W., Fung, P.C.W. and Chow, W.K. (2019) *Journal of Modern Physics*, **10**, 1548-1565. https://doi.org/10.4236/jmp.2019.1013103

[13] Frohlich, H. (1968) *International Journal of Quantum Chemistry*, **2**, 641. https://doi.org/10.1002/qua.560020505

[14] Geesink, H.J. (2020) Proposed Informational Code of Bio Molecules and Its Building Blocks: Quantum Coherence versus Decoherence.

[15] Geesink, H.J.H. and Meijer, D.K.F. (2018) *Journal of Modern Physics*, **9**, 898-924. https://doi.org/10.4236/jmp.2018.95056

[16] Wong, K.W., Dreschhoff, G. and Jungner, H. (2012) The J/Si Meson and the Missing Baryon Octet.

[17] Wong, K.W., Dreschhoff, G. and Jungner, H. (2012) *Journal of Modern Physics*, **3**, 1450-1457. https://doi.org/10.4236/jmp.2012.310179

[18] Einstein, A. (1916) *Annalen der Physik*, **49**, 769. https://doi.org/10.1002/andp.19163540702

[19] Goldstein, A., *et al.* (2017) *Astrophysical Journal Letters*, **848**, L14. https://doi.org/10.3847/2041-8213/aa8f41

[20] Chen, X.D., *et al.* (2019) *Nature Astronomy*, **3**, 320-325. https://doi.org/10.1038/s41550-018-0686-7

[21] Wong, K.W. and Ching, C.W. (1989) *Physica C*, **158**, 1-14. https://doi.org/10.1016/0921-4534(89)90294-3

[22] Ching, W.Y., Xu, Y.N., Zhao, G.L., Wong, K.W. and Zandiehnadem, F. (1987) *Physical Review Letters*, **59**, 1333-1336. https://doi.org/10.1103/PhysRevLett.59.1333

[23] Wong, K.W. and Ching, W.Y. (1989) *Physica C*, **158**, 15-31.

https://doi.org/10.1016/0921-4534(89)90295-5

[24] Wong K.W. and Curatolo, S. (2008) EEM: The Excitonic Enhancement Mechanism Theory and Experimental Evidence of Optically Enhanced in High Tc Superconductors. In: *Progress in Superconductivity Research*, Nova Science Publishers, Inc., Hauppauge, 55-78.

[25] Geesink, H.J.H., Jerman, I. and Meijer, D.K.F. (2020) Clay Minerals Information Network Linking Quantum Coherence versus Decoherence and First Life.

[26] Wong, K.W., Fung, P.C.W., Yeung, H.Y. and Kwok, W.Y. (1992) *Physical Review B*, **45**, 13017-13024. https://doi.org/10.1103/PhysRevB.45.13017

[27] Wong, K.W. and Ching, W.Y. (2004) *Physica C*, **416**, 47-67. https://doi.org/10.1016/j.physc.2004.09.003

[28] McGlinn, W.D. (1964) *Physical Review Letters*, **12**, 467. https://doi.org/10.1103/PhysRevLett.12.467

Journal of Modern Physics, 2020, 11, 1075-1090
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Resolution in the Case of the Hydrogen Atom of an Improved Dirac Equation

Claude Daviau[1], Jacques Bertrand[2], Raymond Albert Ng[3]

[1]Le Moulin de la Lande, Pouillé-les-Côteaux, France
[2]15 Avenue Danielle Casanova, Saint-Gratien, France
[3]De La Salle University (DLSU), Manila, Philippines
Email: daviau.claude@orange.fr, bertrandjacques-m@orange.fr, raymondalbertng@gmail.com

How to cite this paper: Daviau, C., Bertrand, J. and Ng, R.A. (2020) Resolution in the Case of the Hydrogen Atom of an Improved Dirac Equation. *Journal of Modern Physics*, **11**, 1075-1090.
https://doi.org/10.4236/jmp.2020.117068

Received: June 11, 2020
Accepted: July 17, 2020
Published: July 20, 2020

Abstract

The improved Dirac equation is completely solved in the case of the hydrogen atom. A method of separation of variables in spherical coordinates is used. The angular functions are the same as with the linear Dirac equation: they account for the spin 1/2 of the electron. The existence of a probability density governs the radial equations. This gives all the quantum numbers required by spectroscopy, the true number of energy levels and the true levels obtained by Sommerfeld's formula.

Keywords

Electromagnetism, Clifford Algebra, Dirac Equation, Lagrangian Formalism, Hydrogen Atom

1. Introduction

The improved Dirac equation was obtained from Lochak's theory of a leptonic magnetic monopole [1] [2] [3]. In this theory, the quantum wave has two U(1) gauge invariances. The mass term of the leptonic magnetic monopole is then able to replace the usual mass term of the Dirac equation. First, the non-linear mass term was read in the frame of Hestenes's space-time algebra [4] [5] [6]-[11]. The improved equation may be obtained from the simplification of the Lagrangian density of the Dirac equation, where the $m\bar{\psi}\psi = m\cos(\beta)\rho$ mass term is replaced by $m\rho$. The Dirac wave of the linear equation or of the improved equation may be read as functions of space-time with values in the Clifford algebra Cl_3 [12]-[34].

Since 1928, the relativistic invariance of Dirac's theory used the previous Pauli matrices for the spin of the electron: the space-time variable $x = \left(x^0, x^1, x^2, x^3\right)$

was replaced by

$$x = x^0 + \vec{x} = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix},\ x^0 = ct. \tag{1}$$

This is equivalent to saying that the three Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix};\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2}$$

form a orthogonal oriented basis in space. We shall put arrows on vectors in space, so any vector reads

$$\vec{v} = v^j \sigma_j = v^1\sigma_1 + v^2\sigma_2 + v^3\sigma_3. \tag{3}$$

The geometric algebra of space Cl_3 and $M_2(\mathbb{C})$ are isomorphic algebras on the real field; the sum and the product of matrices are familiar in quantum physics. This matrix representation identifies complex numbers and scalar matrices in the Pauli algebra. With this identification, we write the x of (1) as $x = x^\mu \sigma_\mu$, we consider $(\sigma_0, \sigma_1, \sigma_2, \sigma_3)$ as a basis in space-time and we use the Einstein summation convention, with superscripts and subscripts, where Latin indices are in space and Greek indices in space-time. Any element z in the Clifford algebra of space Cl_3 is a sum of a real part x, a vector part $\vec{v}$, an axial-vector part $i\vec{w}$ and a pseudo-scalar part iy, and we need:

$$\begin{aligned} &z = x + \vec{v} + i\vec{w} + iy;\ \hat{z} = x - \vec{v} + i\vec{w} - iy, \\ &\tilde{z} = z^\dagger = x + \vec{v} - i\vec{w} - iy;\ \overline{z} = x - \vec{v} - i\vec{w} + iy. \end{aligned} \tag{4}$$

The application $z \mapsto \hat{z}$ is the main automorphism of Cl_3. The reverse is also the adjoint (transposed conjugate matrix), so $z \mapsto \tilde{z} = z^\dagger$ is the reversion. The third conjugation, $z \mapsto \overline{z}$ is the product of the two previous ones and we shall need:

$$\overline{z} = \hat{z}^\dagger = \widehat{z^\dagger};\ \widehat{AB} = \hat{A}\hat{B};\ \overline{AB} = \overline{B}\,\overline{A};\ M\overline{M} = \overline{M}M = \det(M) \tag{5}$$

Space-time is then made of the auto-adjoint part of the space algebra. We use:

$$\hat{x} = \overline{x} = x^0 - \vec{x};\ x^\dagger = x, \tag{6}$$

$$\det(x) = x\hat{x} = x \cdot x = (x^0)^2 - \vec{x}^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2.$$

The main reason to use the geometric algebra Cl_3 is the ability to read all relativistic quantum physics in this algebra: the electron wave is a function of space and time in Cl_3:

$$x \mapsto \phi = \phi(x) \in Cl_3. \tag{7}$$

The link between Cl_3 and the complex formalism is simple only if we use the left and right Weyl spinors η and ξ, by letting:

$$\begin{aligned} &\phi = \sqrt{2}\begin{pmatrix} \xi & \hat{\eta} \end{pmatrix} = \sqrt{2}\begin{pmatrix} \xi_1 & -\overline{\eta}_2 \\ \xi_2 & \overline{\eta}_1 \end{pmatrix};\ \hat{\eta} = \begin{pmatrix} -\overline{\eta}_2 \\ \overline{\eta}_1 \end{pmatrix}, \\ &\hat{\phi} = \sqrt{2}\begin{pmatrix} \eta & \hat{\xi} \end{pmatrix} = \sqrt{2}\begin{pmatrix} \eta_1 & -\overline{\xi}_2 \\ \eta_2 & \overline{\xi}_1 \end{pmatrix};\ \hat{\xi} = \begin{pmatrix} -\overline{\xi}_2 \\ \overline{\xi}_1 \end{pmatrix}. \end{aligned} \tag{8}$$

The usual formalism uses a ψ and γ^μ matrices defined as following:

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix}; \gamma^0 = \gamma_0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}; I = \sigma^0 = \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{9}$$

$$\gamma^j = -\gamma_j = \begin{pmatrix} 0 & -\sigma_j \\ \sigma_j & 0 \end{pmatrix}, \ j = 1,2,3. \tag{10}$$

Our improved wave equation of the electron, which has the Dirac equation as linear approximation, reads [6]-[19]:

$$\overline{\phi}\left(\nabla\hat{\phi}\right)\sigma_{21} + \overline{\phi}qA\hat{\phi} + m\rho = 0; \nabla = \sigma^\mu\partial_\mu; \sigma_{21} = \sigma_2\sigma_1; \rho = \left|\det\left(\phi\right)\right|. \tag{11}$$

where $q = e/\hbar c$, $m = m_0 c/\hbar$, $\sigma^0 = \sigma_0$, $\sigma^j = -\sigma_j$, $j = 1,2,3$. This equation is invariant under any transformation D defined by an element M of the Lie group Cl_3^* (group of invertible elements of Cl_3):

$$x' = D\left(x\right) = MxM^\dagger = x'^\mu\sigma_\mu; \ \partial'_\mu = \frac{\partial}{\partial x'^\mu}; \ x'^\nu = D^\nu_\mu x^\mu, \tag{12}$$

$$\nabla = \overline{M}\nabla'\hat{M}; \ \nabla' = \sigma^\mu\partial'_\mu; \ qA = \overline{M}q'A'\hat{M}. \tag{13}$$

$$\phi'\left(x'\right) = M\phi\left(x\right); \ \xi' = M\xi; \ \eta' = \hat{M}\eta. \tag{14}$$

Multiplying (11) on the left side by $\overline{\phi}^{-1}$, we obtain the usual form in the Pauli algebra of the improved equation:

$$\nabla\hat{\phi}\sigma_{21} + qA\hat{\phi} + me^{-i\beta}\phi = 0; \ \det\left(\phi\right) = \rho e^{i\beta}, \tag{15}$$

where β is the Yvon-Takabayasi angle. In the usual formalism of complex 4×4 matrices, much more complicated than the Pauli algebra, this wave equation reads:

$$0 = \left[\gamma^\mu\left(\partial_\mu + iqA_\mu\right) + im\exp\left(-\beta\gamma_0\gamma_1\gamma_2\gamma_3\right)\right]\psi. \tag{16}$$

Our wave equation has the Dirac equation as linear approximation: if β is null or negligible, the improved Equation (15) is reduced to the Dirac equation that reads in the Pauli algebra and in the Dirac algebra:

$$0 = \nabla\hat{\phi}\sigma_{21} + qA\hat{\phi} + m\phi, \tag{17}$$

$$0 = \left[\gamma^\mu\left(\partial_\mu + iqA_\mu\right) + im\right]\psi. \tag{18}$$

Our equation is an improvement from the linear Dirac equation for many reasons explained in our previous papers [6]-[34]: Charge conjugation gives a positron with positive mass-energy. The gauge group comprises a local electric gauge linked to the conservative current of probability and a second gauge (only global) that is the chiral gauge of the Yvon-Takabayasi angle, linked to a second conservative current. This $U\left(1\right)$ gauge invariance corresponds to the first part of the $U\left(1\right)\times SU\left(2\right)$ group of electro-weak interactions. Therefore, our improved equation allows us a generalization, with mass term, to all fermion waves of leptons and quarks. In this paper, we shall show how this improved, non-linear wave equation is nevertheless able to obtain the following results: the

true quantum numbers, the true number of energy levels and the true energy levels in the case of the hydrogen atom. This was previously never obtained from a non-linear wave equation; therefore, everyone was thinking that the linearity of quantum mechanics was necessary, even for the linking to the non-linear relativistic gravitation.

2. Separating Variables

To solve the Dirac equation or the improved equation in the case of the hydrogen atom, two methods exist. We shall use here, not the initial method based on the non-relativistic approximation of the wave equation, but the new method invented by H. Krüger [35], separating the variables in spherical coordinates:

$$x^1 = r\sin\theta\cos\varphi;\ x^2 = r\sin\theta\sin\varphi;\ x^3 = r\cos\theta. \tag{19}$$

We use the following notation:

$$i_1 = \sigma_{23} = i\sigma_1;\ i_2 = \sigma_{31} = i\sigma_2;\ i_3 = \sigma_{12} = i\sigma_3, \tag{20}$$

$$S = \mathrm{e}^{-\frac{\phi}{2}i_3}\mathrm{e}^{-\frac{\theta}{2}i_2};\ \Omega = \hat{\Omega} = r^{-1}\left(\sin\theta\right)^{-\frac{1}{2}}S, \tag{21}$$

$$\vec{\partial} = \sigma_1\partial_1 + \sigma_2\partial_2 + \sigma_3\partial_3;\ \vec{\partial}' = \sigma_3\partial_r + \frac{1}{r}\sigma_1\partial_\theta + \frac{1}{r\sin\theta}\sigma_2\partial_\varphi. \tag{22}$$

H. Krüger obtained the remarkable identity:

$$\vec{\partial} = \Omega\vec{\partial}'\Omega^{-1}, \tag{23}$$

which with

$$\nabla' = \partial_0 - \vec{\partial}' = \partial_0 - \left(\sigma_3\partial_r + \frac{1}{r}\sigma_1\partial_\theta + \frac{1}{r\sin\theta}\sigma_2\partial_\varphi\right), \tag{24}$$

gives also:

$$\Omega^{-1}\nabla = \nabla'\Omega^{-1}. \tag{25}$$

Aiming for the separation of the temporal variable $x^0 = ct$ and the angular variable φ from the radial variable r and the other angular variable θ, we let:

$$\phi = \Omega X\mathrm{e}^{\left(\lambda\varphi - Ex^0 + \delta\right)i_3};\ X = \left(\xi\ \hat{\eta}\right), \tag{26}$$

where X is a function (with value in the Pauli algebra) of only r and θ, $\hbar cE$ is the energy of the electron, δ is an arbitrary phase (that plays no role here because the wave equations are electric gauge-invariant), and λ is a real constant which will be interpreted as the magnetic quantum number (we name here this quantum number λ because m is used in the mass term). We then get:

$$\Omega^{-1}\phi = X\mathrm{e}^{\left(\lambda\varphi - Ex^0 + \delta\right)i_3}, \tag{27}$$

$$\Omega^{-1}\hat{\phi} = \hat{X}\mathrm{e}^{\left(\lambda\varphi - Ex^0 + \delta\right)i_3}. \tag{28}$$

We also have:

$$\rho\mathrm{e}^{i\beta} = \det\left(\phi\right) = \det\left(\Omega\right)\det\left(X\right)\det\left[\mathrm{e}^{\left(\lambda\varphi - Ex^0 + \delta\right)i_3}\right], \tag{29}$$

$$\det\left(\Omega\right)=r^{-2}\left(\sin\theta\right)^{-1};\ \det\left[\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}\right]=1, \tag{30}$$

$$\rho\mathrm{e}^{i\beta}=\frac{\det\left(X\right)}{r^2\sin\theta}. \tag{31}$$

Then if we let

$$\rho_X\mathrm{e}^{i\beta_X}=\det\left(X\right), \tag{32}$$

we get:

$$\rho=\frac{\rho_X}{r^2\sin\theta};\ \beta=\beta_X. \tag{33}$$

Then with the form (26) of the wave, the Yvon-Takabayasi angle β depends neither on the time nor on the φ angle. It depends only on r and θ. Hence the separation of variables can be similarly obtained for the Dirac equation or for the improved equation. We have:

$$\nabla'\Omega^{-1}\hat{\phi}=\left(\partial_0-\sigma_3\partial_r-\frac{1}{r}\sigma_1\partial_\theta-\frac{1}{r\sin\theta}\sigma_2\partial_\varphi\right)\left[\hat{X}\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}\right], \tag{34}$$

$$\partial_0\left(\hat{X}\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}\right)=-E\hat{X}i_3\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}, \tag{35}$$

$$\partial_r\left(\hat{X}\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}\right)=\left(\partial_r\hat{X}\right)\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3} \tag{36}$$

$$\partial_\theta\left(\hat{X}\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}\right)=\left(\partial_\theta\hat{X}\right)\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}, \tag{37}$$

$$\partial_\varphi\left(\hat{X}\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}\right)=\lambda\hat{X}i_3\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}. \tag{38}$$

We then get:

$$\nabla\hat{\phi}=\Omega\left(-E\hat{X}i_3-\sigma_3\partial_r\hat{X}-\frac{1}{r}\sigma_1\partial_\theta\hat{X}-\frac{\lambda}{r\sin\theta}\sigma_2\hat{X}i_3\right)\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}. \tag{39}$$

For the hydrogen atom we have:

$$qA=qA^0=-\frac{\alpha}{r};\ \alpha=\frac{\mathrm{e}^2}{\hbar c}, \tag{40}$$

where α is the fine structure constant. We have:

$$qA\hat{\phi}\sigma_{12}=-\frac{\alpha}{r}\hat{\phi}i_3=-\frac{\alpha}{r}\Omega\hat{X}\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}i_3=\Omega\left(-\frac{\alpha}{r}\hat{X}i_3\right)\mathrm{e}^{\left(\lambda\varphi-Ex^0+\delta\right)i_3}. \tag{41}$$

Also the wave Equation (15) becomes:

$$-E\hat{X}i_3-\sigma_3\partial_r\hat{X}-\frac{1}{r}\sigma_1\partial_\theta\hat{X}-\frac{\lambda}{r\sin\theta}\sigma_2\hat{X}i_3-\frac{\alpha}{r}\hat{X}i_3+m\mathrm{e}^{-i\beta}Xi_3=0, \tag{42}$$

which means:

$$\left(E+\frac{\alpha}{r}\right)\hat{X}i_3+\sigma_3\partial_r\hat{X}+\frac{1}{r}\sigma_1\partial_\theta\hat{X}+\frac{\lambda}{r\sin\theta}\sigma_2\hat{X}i_3=m\mathrm{e}^{-i\beta}Xi_3, \tag{43}$$

The Dirac equation, in contrast, gives:

$$\left(E+\frac{\alpha}{r}\right)\hat{X}i_3+\sigma_3\partial_r\hat{X}+\frac{1}{r}\sigma_1\partial_\theta\hat{X}+\frac{\lambda}{r\sin\theta}\sigma_2\hat{X}i_3=mXi_3. \tag{44}$$

Now we let, $\overline{z}$ being the complex conjugate of z:

$$X=\begin{pmatrix}\mathbf{a} & -\overline{\mathbf{b}}\\ \mathbf{c} & \overline{\mathbf{d}}\end{pmatrix}, \tag{45}$$

where $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$ are functions with complex values of the real variables r and θ. We get:

$$\hat{X}=\begin{pmatrix}\mathbf{d} & -\overline{\mathbf{c}}\\ \mathbf{b} & \overline{\mathbf{a}}.\end{pmatrix}. \tag{46}$$

We then obtain the following equations:

$$ime^{-i\beta}X\sigma_3=ie^{-i\beta}\begin{pmatrix}m\mathbf{a} & m\overline{\mathbf{b}}\\ m\mathbf{c} & -m\overline{\mathbf{d}}\end{pmatrix}, \tag{47}$$

$$\hat{X}i_3=\begin{pmatrix}\mathbf{d} & -\overline{\mathbf{c}}\\ \mathbf{b} & \overline{\mathbf{a}}\end{pmatrix}\begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}=\begin{pmatrix}i\mathbf{d} & i\overline{\mathbf{c}}\\ i\mathbf{b} & -i\overline{\mathbf{a}}\end{pmatrix}, \tag{48}$$

$$\sigma_3\partial_r\hat{X}=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\partial_r\mathbf{d} & -\partial_r\overline{\mathbf{c}}\\ \partial_r\mathbf{b} & \partial_r\overline{\mathbf{a}}\end{pmatrix}=\begin{pmatrix}\partial_r\mathbf{d} & -\partial_r\overline{\mathbf{c}}\\ -\partial_r\mathbf{b} & -\partial_r\overline{\mathbf{a}}\end{pmatrix}, \tag{49}$$

$$\sigma_1\partial_\theta\hat{X}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\partial_\theta\mathbf{d} & -\partial_\theta\overline{\mathbf{c}}\\ \partial_\theta\mathbf{b} & \partial_\theta\overline{\mathbf{a}}\end{pmatrix}=\begin{pmatrix}\partial_\theta\mathbf{b} & \partial_\theta\overline{\mathbf{a}}\\ \partial_\theta\mathbf{d} & -\partial_\theta\overline{\mathbf{c}}\end{pmatrix}, \tag{50}$$

$$\sigma_2\hat{X}i_3=i_2\hat{X}\sigma_3=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\mathbf{d} & -\overline{\mathbf{c}}\\ \mathbf{b} & \overline{\mathbf{a}}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}=\begin{pmatrix}\mathbf{b} & -\overline{\mathbf{a}}\\ -\mathbf{d} & -\overline{\mathbf{c}}\end{pmatrix}. \tag{51}$$

Therefore, the improved equation is equivalent to:

$$\begin{aligned}&\left(E+\frac{\alpha}{r}\right)\begin{pmatrix}i\mathbf{d} & i\overline{\mathbf{c}}\\ i\mathbf{b} & -i\overline{\mathbf{a}}\end{pmatrix}+\begin{pmatrix}\partial_r\mathbf{d} & -\partial_r\overline{\mathbf{c}}\\ -\partial_r\mathbf{b} & -\partial_r\overline{\mathbf{a}}\end{pmatrix}+\frac{1}{r}\begin{pmatrix}\partial_\theta\mathbf{b} & \partial_\theta\overline{\mathbf{a}}\\ \partial_\theta\mathbf{d} & -\partial_\theta\overline{\mathbf{c}}\end{pmatrix}\\ &+\frac{\lambda}{r\sin\theta}\begin{pmatrix}\mathbf{b} & -\overline{\mathbf{a}}\\ -\mathbf{d} & -\overline{\mathbf{c}}\end{pmatrix}=ie^{-i\beta}\begin{pmatrix}m\mathbf{a} & m\overline{\mathbf{b}}\\ m\mathbf{c} & -m\overline{\mathbf{d}}\end{pmatrix}.\end{aligned} \tag{52}$$

Conjugating the equations containing the terms of the right columns we get the system:

$$\begin{aligned}&i\left(E+\frac{\alpha}{r}\right)\mathbf{d}+\partial_r\mathbf{d}+\frac{1}{r}\left(\partial_\theta+\frac{\lambda}{\sin\theta}\right)\mathbf{b}=ime^{-i\beta}\mathbf{a},\\ &-i\left(E+\frac{\alpha}{r}\right)\mathbf{c}-\partial_r\mathbf{c}+\frac{1}{r}\left(\partial_\theta-\frac{\lambda}{\sin\theta}\right)\mathbf{a}=-ime^{i\beta}\mathbf{b},\\ &i\left(E+\frac{\alpha}{r}\right)\mathbf{b}-\partial_r\mathbf{b}+\frac{1}{r}\left(\partial_\theta-\frac{\lambda}{\sin\theta}\right)\mathbf{d}=ime^{-i\beta}\mathbf{c},\\ &-i\left(E+\frac{\alpha}{r}\right)\mathbf{a}+\partial_r\mathbf{a}+\frac{1}{r}\left(\partial_\theta+\frac{\lambda}{\sin\theta}\right)\mathbf{c}=-ime^{i\beta}\mathbf{d}.\end{aligned} \tag{53}$$

Moreover we have:

$$\rho e^{i\beta}=\det(\phi)=\frac{\det(X)}{r^2\sin\theta}=\frac{\mathbf{a}\overline{\mathbf{d}}+\mathbf{c}\overline{\mathbf{b}}}{r^2\sin\theta} \tag{54}$$

then we get:

$$e^{i\beta}=\frac{\mathbf{a}\bar{\mathbf{d}}+\mathbf{c}\bar{\mathbf{b}}}{\left|\mathbf{a}\bar{\mathbf{d}}+\mathbf{c}\bar{\mathbf{b}}\right|}. \tag{55}$$

In the Equation (53) only two angular operators are present, then we let:

$$\mathbf{a}=AU;\ \mathbf{b}=BV;\ \mathbf{c}=CV;\ \mathbf{d}=DU, \tag{56}$$

where A, B, C et D are functions of r while U and V are functions of θ. The (53) system becomes:

$$\begin{aligned}
&i\left(E+\frac{\alpha}{r}\right)DU+D'U+\frac{1}{r}\left(V'+\frac{\lambda}{\sin\theta}V\right)B=ime^{-i\beta}AU,\\
&-i\left(E+\frac{\alpha}{r}\right)CV-C'V+\frac{1}{r}\left(U'-\frac{\lambda}{\sin\theta}U\right)A=-ime^{i\beta}BV,\\
&i\left(E+\frac{\alpha}{r}\right)BV-B'V+\frac{1}{r}\left(U'-\frac{\lambda}{\sin\theta}U\right)D=ime^{-i\beta}CV,\\
&-i\left(E+\frac{\alpha}{r}\right)AU+A'U+\frac{1}{r}\left(V'+\frac{\lambda}{\sin\theta}V\right)C=-ime^{i\beta}DU.
\end{aligned} \tag{57}$$

Then if a κ constant exists such that:

$$U'-\frac{\lambda}{\sin\theta}U=-\kappa V;\ V'+\frac{\lambda}{\sin\theta}V=\kappa U, \tag{58}$$

the (57) system becomes:

$$\begin{aligned}
&i\left(E+\frac{\alpha}{r}\right)D+D'+\frac{\kappa}{r}B=ime^{-i\beta}A\\
&-i\left(E+\frac{\alpha}{r}\right)C-C'-\frac{\kappa}{r}A=-ime^{i\beta}B\\
&i\left(E+\frac{\alpha}{r}\right)B-B'-\frac{\kappa}{r}D=ime^{-i\beta}C\\
&-i\left(E+\frac{\alpha}{r}\right)A+A'+\frac{\kappa}{r}C=-ime^{i\beta}D
\end{aligned} \tag{59}$$

To get the system equivalent to the Dirac equation it is enough to suppress the β angle. This does not change the angular system (58), while in the place of (59) we get the system:

$$\begin{aligned}
&i\left(E+\frac{\alpha}{r}\right)D+D'+\frac{\kappa}{r}B=imA,\\
&-i\left(E+\frac{\alpha}{r}\right)C-C'-\frac{\kappa}{r}A=-imB,\\
&i\left(E+\frac{\alpha}{r}\right)B-B'-\frac{\kappa}{r}D=imC,\\
&-i\left(E+\frac{\alpha}{r}\right)A+A'+\frac{\kappa}{r}C=-imD.
\end{aligned} \tag{60}$$

3. Kinetic Momentum Operators

We established in [8] the form that the operators of kinetic momentum have in space-time. With the Pauli algebra we have (a detailed calculation is in [13] A.3):

$$J_1\phi=\left(d_1+\frac{1}{2}\sigma_{23}\right)\phi\sigma_{21};\ d_1=x^2\partial_3-x^3\partial_2=-\sin\varphi\,\partial_\theta-\frac{\cos\varphi}{\tan\theta}\partial_\varphi, \tag{61}$$

$$J_2\phi=\left(d_2+\frac{1}{2}\sigma_{31}\right)\phi\sigma_{21};\ d_2=x^3\partial_1-x^1\partial_3=\cos\varphi\,\partial_\theta-\frac{\sin\varphi}{\tan\theta}\partial_\varphi, \tag{62}$$

$$J_3\phi=\left(d_3+\frac{1}{2}\sigma_{12}\right)\phi\sigma_{21};\ d_3=x^1\partial_2-x^2\partial_1=\partial_\varphi. \tag{63}$$

We indeed have also:

$$J^2=J_1^2+J_2^2+J_3^2. \tag{64}$$

From (26) we have the following equivalence for J_3:

$$J_3\phi=\lambda\phi\Leftrightarrow\phi=\phi\left(x^0,r,\theta\right)\mathrm{e}^{\lambda\varphi i_3}. \tag{65}$$

Then the ϕ wave satisfying (26) is a proper vector of J_3 and λ is the magnetic quantum number. Moreover for a wave ϕ satisfying (26), we have:

$$J^2\phi=j\left(j+1\right)\phi, \tag{66}$$

if and only if:

$$\partial_{\theta\theta}^2X+\left[\left(j+\frac{1}{2}\right)^2-\frac{\lambda^2}{\sin^2\theta}\right]X-\lambda\frac{\cos\theta}{\sin^2\theta}\sigma_{12}X\sigma_{12}=0. \tag{67}$$

And (58) implies at the second order:

$$0=U''+\left(\kappa^2-\frac{\lambda^2}{\sin^2\theta}\right)U+\lambda\frac{\cos\theta}{\sin^2\theta}U \tag{68}$$

$$0=V''+\left(\kappa^2-\frac{\lambda^2}{\sin^2\theta}\right)V-\lambda\frac{\cos\theta}{\sin^2\theta}V \tag{69}$$

$$0=\partial_{\theta\theta}^2X+\left(\kappa^2-\frac{\lambda^2}{\sin^2\theta}\right)X-\lambda\frac{\cos\theta}{\sin^2\theta}\sigma_{12}X\sigma_{12}, \tag{70}$$

therefore ϕ is a proper vector of J^2, with the proper value $j\left(j+1\right)$, if and only if:

$$\kappa^2=\left(j+\frac{1}{2}\right)^2;\ \left|\kappa\right|=j+\frac{1}{2};\ j=\left|\kappa\right|-\frac{1}{2}. \tag{71}$$

with the definition of $\mathcal{S}$ in (21) and with (26) we can see that the change of φ into $\varphi+2\pi$ conserves the same value for the wave if and only if λ has a half-integer value. It is only in this case that the wave is correctly defined. The general results on the angular momentum operators imply then:

$$j=\frac{1}{2},\frac{3}{2},\frac{5}{2},\cdots;\ \kappa=\pm1,\pm2,\pm3,\cdots;\ \lambda=-j,-j+1,\cdots j-1,\ j. \tag{72}$$

To solve the angular system we let, if $\lambda>0$ and with $C=C\left(\theta\right)$:

$$\begin{aligned}U&=\sin^\lambda\theta\left[\sin\left(\frac{\theta}{2}\right)C'-\left(\kappa+\frac{1}{2}-\lambda\right)\cos\left(\frac{\theta}{2}\right)C\right],\\ V&=\sin^\lambda\theta\left[\cos\left(\frac{\theta}{2}\right)C'+\left(\kappa+\frac{1}{2}-\lambda\right)\sin\left(\frac{\theta}{2}\right)C\right].\end{aligned} \tag{73}$$

while if $\lambda < 0$ we let:

$$\begin{aligned} U &= \sin^{-\lambda}\theta\left[\cos\left(\frac{\theta}{2}\right)C' + \left(\kappa + \frac{1}{2} + \lambda\right)\sin\left(\frac{\theta}{2}\right)C\right], \\ V &= \sin^{-\lambda}\theta\left[-\sin\left(\frac{\theta}{2}\right)C' + \left(\kappa + \frac{1}{2} + \lambda\right)\cos\left(\frac{\theta}{2}\right)C\right]. \end{aligned} \tag{74}$$

The angular system (58) is then equivalent [6] to the differential equation:

$$0 = C'' + \frac{2|\lambda|}{\tan\theta}C' + \left[\left(\kappa + \frac{1}{2}\right)^2 - \lambda^2\right]C. \tag{75}$$

The change of variable:

$$z = \cos\theta;\ \ f(z) = C\left[\theta(z)\right], \tag{76}$$

gives then the differential equation of the Gegenbauer polynomials[1]:

$$0 = f''(z) - \frac{1 + 2|\lambda|}{1 - z^2}zf'(z) + \frac{\left(\kappa + \frac{1}{2}\right)^2 - \lambda^2}{1 - z^2}f(z). \tag{77}$$

And we get, as only integrable function:

$$\frac{C(\theta)}{C(0)} = \sum_{n=0}^{\infty}\frac{\left(|\lambda| - \kappa - \frac{1}{2}\right)_n\left(|\lambda| + \kappa + \frac{1}{2}\right)_n}{\left(\frac{1}{2} + |\lambda|\right)_n n!}\sin^{2n}\left(\frac{\theta}{2}\right), \tag{78}$$

with:

$$(a)_0 = 1,\quad (a)_n = a(a+1)\cdots(a+n-1). \tag{79}$$

The $C(0)$ term is a factor of U and V, its argument may be absorbed by the δ of (26), and its modulus may be transferred to the radial functions. We can then let $C(0) = 1$, which gives:

$$C(\theta) = \sum_{n=0}^{\infty}\frac{\left(|\lambda| - \kappa - \frac{1}{2}\right)_n\left(|\lambda| + \kappa + \frac{1}{2}\right)_n}{\left(\frac{1}{2} + |\lambda|\right)_n n!}\sin^{2n}\left(\frac{\theta}{2}\right). \tag{80}$$

Since we have the (71) conditions on λ and κ, an integer n always exists such as:

$$|\lambda| + n = \left|\kappa + \frac{1}{2}\right|. \tag{81}$$

This constrains the (80) series to be a finite sum, so then U and V are integrable. And since U and V have real values, we have:

$$\mathrm{e}^{i\beta} = \frac{A\bar{D}U^2 + C\bar{B}V^2}{\left|A\bar{D}U^2 + C\bar{B}V^2\right|}. \tag{82}$$

[1]When we solve the Dirac equation with Darwin's method, that means with the ad-hoc operators, we get some Legendre polynomials and spherical harmonics. Here, working with ϕ, which is equivalent to working with the Weyl spinors ξ and η, we get the Gegenbauer polynomials, and it is the degree of these polynomials that gives the needed quantum number.

4. Resolution of the Linear Radial System

We employ the following transformations:

$$\begin{aligned} &x = mr;\ \epsilon = \frac{E}{m};\ a(x) = A(r) = A\left(\frac{x}{m}\right), \\ &b(x) = B(r);\ c(x) = C(r);\ d(x) = D(r). \end{aligned} \tag{83}$$

The (60) system becomes:

$$\begin{aligned} &i\left(\epsilon + \frac{\alpha}{x}\right)d + d' + \frac{\kappa}{x}b = ia, \\ &-i\left(\epsilon + \frac{\alpha}{x}\right)c - c' - \frac{\kappa}{x}a = -ib, \\ &i\left(\epsilon + \frac{\alpha}{x}\right)b - b' - \frac{\kappa}{x}d = ic, \\ &-i\left(\epsilon + \frac{\alpha}{x}\right)a + a' + \frac{\kappa}{x}c = -id. \end{aligned} \tag{84}$$

And the (59) system becomes:

$$i\left(\epsilon + \frac{\alpha}{x}\right)d + d' + \frac{\kappa}{x}b = ie^{-i\beta}a, \tag{85}$$

$$-i\left(\epsilon + \frac{\alpha}{x}\right)c - c' - \frac{\kappa}{x}a = -ibe^{i\beta}, \tag{86}$$

$$i\left(\epsilon + \frac{\alpha}{x}\right)b - b' - \frac{\kappa}{x}d = ice^{-i\beta}, \tag{87}$$

$$-i\left(\epsilon + \frac{\alpha}{x}\right)a + a' + \frac{\kappa}{x}c = -ide^{i\beta}. \tag{88}$$

Since this radial system has the same asymptotic behavior as its linear approximation (84) we are left with the following as the only integrable solution:

$$\begin{aligned} &a = e^{-\Lambda x}x^s\left(a_0 + a_1x + \cdots + a_nx^n\right) \\ &b = e^{-\Lambda x}x^s\left(b_0 + b_1x + \cdots + b_nx^n\right) \\ &c = e^{-\Lambda x}x^s\left(c_0 + c_1x + \cdots + c_nx^n\right) \\ &d = e^{-\Lambda x}x^s\left(d_0 + d_1x + \cdots + d_nx^n\right);\ \Lambda = \sqrt{1-\epsilon^2}. \end{aligned} \tag{89}$$

We now study the case where $n > 0$. The (88) equation is equivalent to:

$$\begin{aligned} &i\epsilon\left(\ + a_0x + \cdots + a_{n-1}x^n + a_nx^{n+1}\right) + i\alpha\left(a_0 + a_1x + \cdots + a_nx^n\right) \\ &+ \Lambda\left(\ + a_0x + \cdots + a_{n-1}x^n + a_nx^{n+1}\right) - s\left(a_0 + a_1x + \cdots + a_nx^n\right) \\ &-\left(\ + a_1x + \cdots + na_nx^n\right) - \kappa\left(c_0 + c_1x + \cdots + c_nx^n\right) \\ &= ie^{i\beta}\left(\ + d_0x + \cdots + d_{n-1}x^n + d_nx^{n+1}\right). \end{aligned} \tag{90}$$

Similarly, Equation (87) is equivalent to:

$$\begin{aligned} &i\epsilon\left(\ + b_0x + \cdots + b_{n-1}x^n + b_nx^{n+1}\right) + i\alpha\left(b_0 + b_1x + \cdots + b_nx^n\right) \\ &+ \Lambda\left(\ + b_0x + \cdots + b_{n-1}x^n + b_nx^{n+1}\right) - s\left(b_0 + b_1x + \cdots + b_nx^n\right) \end{aligned}$$

$$\begin{aligned}&-\left(\ +b_1x+\cdots+nb_nx^n\right)-\kappa\left(d_0+d_1x+\cdots+d_nx^n\right)\\&=i\mathrm{e}^{-i\beta}\left(\ +c_0x+\ldots+c_{n-1}x^n+c_nx^{n+1}\right).\end{aligned}\tag{91}$$

Next, Equation (86) is equivalent to:

$$\begin{aligned}&i\epsilon\left(\ +c_0x+\cdots+c_{n-1}x^n+c_nx^{n+1}\right)+i\alpha\left(c_0+c_1x+\cdots+c_nx^n\right)\\&-\Lambda\left(\ +c_0x+\cdots+c_{n-1}x^n+c_nx^{n+1}\right)+s\left(c_0+c_1x+\cdots+c_nx^n\right)\\&+\left(\ +c_1x+\cdots+nc_nx^n\right)+\kappa\left(a_0+a_1x+\cdots+a_nx^n\right)\\&=i\mathrm{e}^{i\beta}\left(\ +b_0x+\cdots+b_{n-1}x^n+b_nx^{n+1}\right).\end{aligned}\tag{92}$$

Finally, Equation (85) is equivalent to:

$$\begin{aligned}&i\epsilon\left(\ +d_0x+\cdots+d_{n-1}x^n+d_nx^{n+1}\right)+i\alpha\left(d_0+d_1x+\cdots+d_nx^n\right)\\&-\Lambda\left(\ +d_0x+\cdots+d_{n-1}x^n+d_nx^{n+1}\right)+s\left(d_0+d_1x+\cdots+d_nx^n\right)\\&+\left(\ +d_1x+\cdots+nd_nx^n\right)+\kappa\left(b_0+b_1x+\cdots+b_nx^n\right)\\&=i\mathrm{e}^{-i\beta}\left(\ +a_0x+\cdots+a_{n-1}x^n+a_nx^{n+1}\right).\end{aligned}\tag{93}$$

We then arrive at three kinds of systems: index 0, index between 0 and n, and index n. For the null index the system is independent of β:

$$\begin{aligned}&\left(i\alpha-s\right)a_0-\kappa c_0=0;\ \left(i\alpha-s\right)b_0-\kappa d_0=0,\\&\kappa a_0+\left(i\alpha+s\right)c_0=0;\ \kappa b_0+\left(i\alpha+s\right)d_0=0.\end{aligned}\tag{94}$$

This system is the same as in the linear case. We obtain a non-null solution only if the determinant is null, then only if s satisfies:

$$0=\kappa^2+\left(i\alpha+s\right)\left(i\alpha-s\right);\ s=\sqrt{\kappa^2-\alpha^2}.\tag{95}$$

And the system (94) is then equivalent to:

$$c_0=\frac{i\alpha-s}{\kappa}a_0;\ d_0=\frac{i\alpha-s}{\kappa}b_0.\tag{96}$$

For the n index the resulting system is:

$$\begin{aligned}&\left(i\epsilon+\Lambda\right)a_n=i\mathrm{e}^{i\beta}d_n;\ \left(i\epsilon+\Lambda\right)b_n=i\mathrm{e}^{-i\beta}c_n;\\&\left(i\epsilon-\Lambda\right)d_n=i\mathrm{e}^{-i\beta}a_n;\ \left(i\epsilon-\Lambda\right)c_n=i\mathrm{e}^{i\beta}b_n.\end{aligned}\tag{97}$$

From the definition of Λ in (89) we obtain the cancellation of the determinants in the two previous sub-systems, and the (97) system is equivalent to:

$$\begin{aligned}&d_n=\left(\epsilon-i\Lambda\right)\mathrm{e}^{-i\beta}a_n,\\&c_n=\left(\epsilon-i\Lambda\right)\mathrm{e}^{i\beta}b_n.\end{aligned}\tag{98}$$

Next the system with index between 0 and n reads:

$$\left(i\epsilon+\Lambda\right)a_{n-1}+\left(i\alpha-s-n\right)a_n-\kappa c_n=i\mathrm{e}^{i\beta}d_{n-1},\tag{99}$$

$$\left(i\epsilon-\Lambda\right)d_{n-1}+\left(i\alpha+s+n\right)d_n+\kappa b_n=i\mathrm{e}^{-i\beta}a_{n-1},\tag{100}$$

$$\left(i\epsilon+\Lambda\right)b_{n-1}+\left(i\alpha-s-n\right)b_n-\kappa d_n=i\mathrm{e}^{-i\beta}c_{n-1},\tag{101}$$

$$\left(i\epsilon-\Lambda\right)c_{n-1}+\left(i\alpha+s+n\right)c_n+\kappa a_n=i\mathrm{e}^{i\beta}b_{n-1}.\tag{102}$$

Now we multiply (99) by $i\epsilon-\Lambda$:

$$(i\epsilon-\Lambda)(i\epsilon+\Lambda)a_{n-1}+(i\epsilon-\Lambda)(i\alpha-s-n)a_n-(i\epsilon-\Lambda)\kappa c_n=ie^{i\beta}(i\epsilon-\Lambda)d_{n-1}. \quad (103)$$

Next (100) implies:

$$(i\epsilon-\Lambda)d_{n-1}=ie^{-i\beta}a_{n-1}-(i\alpha+s+n)d_n-\kappa b_n. \quad (104)$$

And (103) becomes:

$$\begin{aligned}&-a_{n-1}+(i\epsilon-\Lambda)(i\alpha-s-n)a_n-\kappa(i\epsilon-\Lambda)c_n\\&=ie^{i\beta}\left[ie^{-i\beta}a_{n-1}-(i\alpha+s+n)d_n-\kappa b_n\right]\end{aligned}$$

$$(i\epsilon-\Lambda)(i\alpha-s-n)a_n-\kappa(i\epsilon-\Lambda)c_n=-ie^{i\beta}(i\alpha+s+n)d_n-i\kappa e^{i\beta}b_n. \quad (105)$$

Using (97) we get:

$$\begin{aligned}e^{i\beta}d_n&=(\epsilon-i\Lambda)a_n,\\e^{i\beta}b_n&=(\epsilon+i\Lambda)c_n.\end{aligned} \quad (106)$$

Substituting these relations into (105) we obtain:

$$\begin{aligned}&(\epsilon+i\Lambda)\left[-\alpha-i(s+n)\right]a_n-i\kappa(\epsilon+i\Lambda)c_n\\&=-i(i\alpha+s+n)(\epsilon-i\Lambda)a_n-i\kappa(\epsilon+i\Lambda)c_n.\end{aligned} \quad (107)$$

Since $a_n\neq 0$ we get:

$$-(\epsilon+i\Lambda)\left[\alpha+i(s+n)\right]=\left[\alpha-i(s+n)\right](\epsilon-i\Lambda), \quad (108)$$

$$i\Lambda\left[\alpha-i(s+n)-\alpha-i(s+n)\right]=\epsilon\left[\alpha-i(s+n)+\alpha+i(s+n)\right], \quad (109)$$

$$\Lambda(s+n)=\epsilon\alpha. \quad (110)$$

Squaring the equation and substituting the definition of Λ we get:

$$(1-\epsilon^2)(s+n)^2=\epsilon^2\alpha^2, \quad (111)$$

$$(s+n)^2=\epsilon^2\left[\alpha^2+(s+n)^2\right], \quad (112)$$

$$\frac{E^2}{m^2}=\frac{(s+n)^2}{\alpha^2+(s+n)^2}, \quad (113)$$

which gives the Sommerfeld energy levels:

$$E=\frac{m}{\sqrt{1+\dfrac{\alpha^2}{(s+n)^2}}};\ s=\sqrt{\kappa^2-\alpha^2};\ |\kappa|=j+\frac{1}{2}\in\mathbb{N}^*. \quad (114)$$

5. Constant Radial Polynomials

To arrive at all the results of the Dirac equation, there is one last thing to explain: why we have $2\mathbf{n}^2$ different bound states with principal quantum number $\mathbf{n}=|\kappa|+n$. For this we must return to the particular case where the radial polynomial is reduced to a constant. We begin directly from (84), and we let:

$$a=a_0e^{-\Lambda x}x^s;\ b=b_0e^{-\Lambda x}x^s;\ c=c_0e^{-\Lambda x}x^s;\ d=d_0e^{-\Lambda x}x^s. \quad (115)$$

We obtain the following set of equations from (59):

$$\begin{aligned}
&\mathrm{e}^{-\Lambda x}\left(i\epsilon d_0 x^s + i\alpha d_0 x^{s-1} - \Lambda d_0 x^s + s d_0 x^{s-1} + \kappa b_0 x^{s-1}\right) = i a_0 \mathrm{e}^{-\Lambda x} x^s,\\
&\mathrm{e}^{-\Lambda x}\left(-i\epsilon c_0 x^s - i\alpha c_0 x^{s-1} + \Lambda c_0 x^s - s c_0 x^{s-1} - \kappa a_0 x^{s-1}\right) = -i b_0 \mathrm{e}^{-\Lambda x} x^s,\\
&\mathrm{e}^{-\Lambda x}\left(i\epsilon b_0 x^s + i\alpha b_0 x^{s-1} + \Lambda b_0 x^s - s b_0 x^{s-1} - \kappa d_0 x^{s-1}\right) = i c_0 \mathrm{e}^{-\Lambda x} x^s,\\
&\mathrm{e}^{-\Lambda x}\left(-i\epsilon a_0 x^s - i\alpha a_0 x^{s-1} - \Lambda a_0 x^s + s a_0 x^{s-1} + \kappa c_0 x^{s-1}\right) = -i d_0 \mathrm{e}^{-\Lambda x} x^s.
\end{aligned} \tag{116}$$

This is equivalent to the set formed by the four following systems:

$$\begin{aligned}
&\kappa b_0 + (i\alpha + s) d_0 = 0,\\
&(i\alpha - s) b_0 - \kappa d_0 = 0,
\end{aligned} \tag{117}$$

$$\begin{aligned}
&-\kappa a_0 - (i\alpha + s) c_0 = 0,\\
&-(i\alpha - s) a_0 + \kappa c_0 = 0,
\end{aligned} \tag{118}$$

$$\begin{aligned}
&-i a_0 + (i\epsilon - \Lambda) d_0 = 0,\\
&-(i\epsilon + \Lambda) a_0 + i d_0 = 0,
\end{aligned} \tag{119}$$

$$\begin{aligned}
&i b_0 - (i\epsilon - \Lambda) c_0 = 0,\\
&(i\epsilon + \Lambda) b_0 - i c_0 = 0.
\end{aligned} \tag{120}$$

The cancellation of the determinant in (117) and (118) gives again (95). The cancellation of the determinant in (117) and (118) is simply equivalent to $\Lambda^2 = 1 - \epsilon^2$, which results from the definition of Λ. Each system (117) to (120) is then reduced to only one equation:

$$\begin{aligned}
&\kappa d_0 = (i\alpha - s) b_0,\\
&\kappa c_0 = (i\alpha - s) a_0,\\
&d_0 = (\epsilon - i\Lambda) a_0,\\
&b_0 = (\epsilon + i\Lambda) c_0.
\end{aligned} \tag{121}$$

We then obtain:

$$\kappa d_0 = \kappa(\epsilon - i\Lambda) a_0 = (i\alpha - s))_0 = (i\alpha - s)(\epsilon + i\Lambda) c_0 = \frac{(i\alpha - s)^2 (\epsilon + i\Lambda)}{\kappa} a_0. \tag{122}$$

We have a non-null solution only if:

$$\begin{aligned}
&\kappa(\epsilon - i\Lambda) = \frac{(i\alpha - s)^2 (\epsilon + i\Lambda)}{\kappa},\\
&\kappa^2 (\epsilon - i\Lambda)^2 = (s - i\alpha)^2,\\
&\kappa(\epsilon - i\Lambda) = \pm(s - i\alpha).
\end{aligned} \tag{123}$$

Since ϵ, s, Λ and α are positive we finally get:

$$|\kappa| = \frac{s}{\epsilon} = \frac{\alpha}{\Lambda}. \tag{124}$$

This equality gives again the formula of the energy levels (114) with $n = 0$. Since κ comes with an absolute value, we can so well have $\kappa < 0$ as $\kappa > 0$. But Darwin's solutions, which work with real constants and not with complex

ones at this stage of the calculation, forbid κ to be negative and it is this interdiction which induces, for a principal quantum number $\mathbf{n} = n + |\kappa|$, a result of $\mathbf{n}(\mathbf{n}+1) + \mathbf{n}(\mathbf{n}-1) = 2\mathbf{n}^2$ states. This is what really happens: The change of sign of κ is equivalent to changing V into $-V$. And if we change the sign of κ and of V, then $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ are invariant if $n = 0$, and the wave remains unchanged. Changing the sign of κ brings no more supplementary solution and we can then use only the solutions with $\kappa > 0$, in the case $n = 0$. And this allows us to get the true number of states.

6. Concluding Remarks

Sommerfeld's formula (114) for energy levels does not account for the Lamb effect which is, if $n > 0$, a very small shift between the energy levels with the same quantum numbers but with opposite values of κ. If the (114) formula was not the same for opposite values of κ, we should not be able to get four polynomial radial functions with only one condition which gives the quantification of the energy levels. The Standard Model has a precise answer using vacuum polarization. But the calculation must be revised, both to avoid divergences and to employ the improved wave equation which accounts also for weak interactions [33] [34]. Since the Lamb shift is of the same order as the hyperfine structure coming from the interaction between the magnetic moment of the proton and that of the electron, a true calculation must account for the origin of the magnetic moment of the proton from the waves of the three quarks inside, and for the true potential seen by the electron wave, with both the electric charges and the potential vector of the moving charges.

Conflicts of Interest

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

References

[1] Lochak, G. (1983) *Annales de la Fondation Louis de Broglie*, **8**, 4.

[2] Lochak, G. (1984) *Annales de la Fondation Louis de Broglie*, **9**, 1.

[3] Lochak, G. (1985) *International Journal of Theoretical Physics*, **24**, 1019-1050. https://doi.org/10.1007/BF00670815

[4] Hestenes, D. (1973) *Journal of Mathematical Physics*, **14**, 893-905. https://doi.org/10.1063/1.1666413

[5] Hestenes, D. (1986) A Unified Language for Mathematics and Physics and Clifford Algebra and the Interpretation of Quantum Mechanics. In: Chisholm and Common, A., Eds., *Clifford Algebras and Their Applications in Mathematics and Physics*, Reidel, Dordrecht, 1-25, 321-346.

[6] Daviau, C. (1993) Equation de Dirac non linéaire. PhD Thesis, Université de Nantes, Nantes.

[7] Daviau, C. (1997) *Advances in Applied Clifford Algebras*, **7**, 175-194.

[8] Daviau, C. (1997) *Annales de la Fondation Louis de Broglie*, **22**, 87-103.

[9] Daviau, C. (1998) *Annales de la Fondation Louis de Broglie*, **23**, 1.

[10] Daviau, C. (2001) *Annales de la Fondation Louis de Broglie*, **26**, 149-171.

[11] Daviau, C. (2005) *Annales de la Fondation Louis de Broglie*, **30**, 3-4.

[12] Daviau, C. (2011) L'espace-temps double. JePublie, Pouillé-les-coteaux.

[13] Daviau, C. (2012) Double Space-Time and More. JePublie, Pouillé-les-coteaux.

[14] Daviau, C. (2012) Nonlinear Dirac Equation, Magnetic Monopoles and Double Space-Time. CISP, Cambridge.

[15] Daviau, C. (2012) *Advances in Applied Clifford Algebras*, **22**, 611-623. https://doi.org/10.1007/s00006-012-0351-7

[16] Daviau, C. and Bertrand, J. (2012) *Annales de la Fondation Louis de Broglie*, **37**, 129-134.

[17] Daviau, C. and Bertrand, J. (2014) New Insights in the Standard Model of Quantum Physics in Clifford Algebra. Je Publie, Pouillé-les-coteaux. http://hal.archives-ouvertes.fr/hal-00907848

[18] Daviau, C. and Bertrand, J. (2015) The Standard Model of Quantum Physics in Clifford Algebra. World Science Publishing, Singapore. https://doi.org/10.1142/9780

[19] Daviau, C. (2015) *Advances in Applied Clifford Algebras*, **27**, 279-290. https://doi.org/10.1007/s00006-015-0566-5

[20] Daviau, C. (2013) Invariant Quantum Wave Equations and Double Space-Time. In: *Advances in Imaging and Electron Physics*, Vol. 179, Elsevier, Amsterdam, 1-137. https://doi.org/10.1016/B978-0-12-407700-3.00001-6

[21] Daviau, C. and Bertrand, J. (2014) *JMP*, **5**, 1001-1022. https://doi.org/10.4236/jmp.2014.511102

[22] Daviau, C. and Bertrand, J. (2014) *JMP*, **5**, 2149-2173. https://doi.org/10.4236/jmp.2014.518210

[23] Daviau, C. and Bertrand, J. (2015) *Journal of Modern Physics*, **6**, 1647-1656. https://doi.org/10.4236/jmp.2015.611166

[24] Daviau, C. and Bertrand, J. (2015) *Journal of Applied Mathematics and Physics*, **3**, 46-61. https://doi.org/10.4236/jamp.2015.31007

[25] Daviau, C. and Bertrand, J. (2015) *Journal of Modern Physics*, **6**, 2080-2092. https://doi.org/10.4236/jmp.2015.614215

[26] Daviau, C. and Bertrand, J. (2015) *Annales de la Fondation Louis de Broglie*, **40**, 181-209.

[27] Daviau, C. and Bertrand, J. (2016) *Journal of Modern Physics*, **7**, 936-951. https://doi.org/10.4236/jmp.2016.79086

[28] Daviau, C., Bertrand, J. and Girardot, D. (2016) *JMP*, **7**, 1568-1590. https://doi.org/10.4236/jmp.2016.712143

[29] Daviau, C., Bertrand, J. and Girardot, D. (2016) *Journal of Modern Physics*, **7**, 2398-2417. https://doi.org/10.4236/jmp.2016.716207

[30] Daviau, C., Bertrand, J., Girardot, D. and Socroun, T. (2017) *Annales de la Fondation Louis de Broglie*, **42**, 351-377.

[31] Daviau, C. and Bertrand, J. (2019) *Journal of Modern Physics*, **9**, 250-258. https://doi.org/10.4236/jmp.2018.92017

[32] Daviau, C. and Bertrand, J. (2019) *Annales de la Fondation Louis de Broglie*, **44**, 163-186.

[33] Daviau, C., Bertrand, J., Socroun, T. and Girardot, D. (2019) Modèle standard et gravitation. Presses des Mines, Paris.

[34] Daviau, C., Bertrand, J., Socroun, T. and Girardot, D. (2020) Developing a Theory of Everything. Annales de la Fondation Louis de Broglie, Paris.

[35] Krüger, H. (1991) New Solutions of the Dirac Equation for Central Fields. In: Hestenes, D. and Weingartshofer, A., Eds., *The Electron*, Kluwer, Dordrecht, 49-81. https://doi.org/10.1007/978-94-011-3570-2_4

Journal of Modern Physics, 2020, 11, 1091-1122
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Unifying Theory of Dark Energy, Dark Matter, and Baryonic Matter in the Positive-Negative Mass Universe Pair: Protogalaxy and Galaxy Evolutions

Ding-Yu Chung

Utica, Michigan, USA
Email: dy_chung@yahoo.com

How to cite this paper: Chung, D.-Y. (2020) A Unifying Theory of Dark Energy, Dark Matter, and Baryonic Matter in the Positive-Negative Mass Universe Pair: Protogalaxy and Galaxy Evolutions. *Journal of Modern Physics*, **11**, 1091-1122.
https://doi.org/10.4236/jmp.2020.117069

Received: June 21, 2020
Accepted: July 20, 2020
Published: July 23, 2020

Abstract

This paper modifies the Farnes' unifying theory of dark energy and dark matter which are negative-mass, created continuously from the negative-mass universe in the positive-negative mass universe pair. The first modification explains that observed dark energy is 68.6%, greater than 50% for the symmetrical positive-negative mass universe pair. This paper starts with the proposed positive-negative-mass 11D universe pair (without kinetic energy) which is transformed into the positive-negative mass 10D universe pair and the external dual gravities as in the Randall-Sundrum model, resulting in the four equal and separate universes consisting of the positive-mass 10D universe, the positive-mass massive external gravity, the negative-mass 10D universe and the negative-mass massive external gravity. The positive-mass 10D universe is transformed into 4D universe (home universe) with kinetic energy through the inflation and the Big Bang to create positive-mass dark matter which is five times of positive-mass baryonic matter. The other three universes without kinetic energy oscillate between 10D and 10D through 4D, resulting in the hidden universes when D > 4 and dark energy when D = 4, which is created continuously to our 4D home universe with the maximum dark energy = 3/4 = 75%. In the second modification to explain dark matter in the CMB, dark matter initially is not repulsive. The condensed baryonic gas at the critical surface density induces dark matter repulsive force to transform dark matter in the region into repulsive dark matter repulsing one another. The calculated percentages of dark energy, dark matter, and baryonic matter are 68.6 (as an input from the observation), 26 and 5.2, respectively, in agreement with observed 68.6, 26.5 and 4.9, respectively, and dark energy started in 4.33 billion years ago in agreement with the observed 4.71 ±

0.98 billion years ago. In conclusion, the modified Farnes' unifying theory reinterprets the Farnes' equations, and is a unifying theory of dark energy, dark matter, and baryonic matter in the positive-negative mass universe pair. The unifying theory explains protogalaxy and galaxy evolutions in agreement with the observations.

Keywords

Unifying Theory, Farnes, Dark Energy, Dark Matter, Baryonic Matter, Negative Mass, Positive-Negative Mass Universe Pair, Protogalaxy Evolution, Galaxy Evolution

1. Introduction

In the Farnes' unifying theory of dark energy and dark matter [1], dark energy and dark matter can be unified into a single negative-mass dark fluid which is created continuously from the negative-mass universe to our positive-mass universe. The model is a modified ΛCDM cosmology, and indicates that continuously-created negative masses can resemble the cosmological constant and can flatten the rotation curves of galaxies. In the first three-dimensional N-body simulations of negative mass dark matter in the scientific literature, dark matter naturally forms halos around galaxies that extend to several galactic radii. These halos are not cuspy, unlike the cuspy halos derived from the simulations of conventional positive-mass dark matter [2]. The core cusp is not observed [3]. Therefore, the Farnes' cosmological model is able to predict correctly the observed non-cuspy distribution of dark matter in galaxies from first principles.

This paper modifies and reinterprets the Farnes' unifying theory of dark energy and dark matter in the positive-negative mass universe pair model where positive-mass and negative-mass universes are symmetrical. (The Farnes' unifying theory does not assume symmetrical positive-negative mass universes.) The first modification is about dark energy. With the symmetry between positive-mass universe and negative-mass universe, the portion of dark energy to be created continuously from the negative-mass universe to our positive-mass universe cannot start from zero to more than 50%. Since the observed portion of dark energy is 68.6% [4], for symmetrical positive-negative mass universe pair, dark energy is more than negative mass from the symmetrical negative-mass universe. As a result, the source of dark energy needs modification. Another modification is about dark matter. Evidence for early existence of dark matter comes from measurements on cosmological scales of anisotropies in the cosmic microwave background [5] [6] [7]. The CMB is the remnant radiation from the hot early days of the universe. The photons underwent oscillations that froze in just before decoupling from the baryonic matter at a redshift of 1100. The angular scale and height of the peaks (and troughs) of these oscillations are powerful probes of cosmological parameters, including the total energy density, the ba-

ryonic fraction, and the dark matter component. The CMB provides irrefutable evidence for the early existence of dark matter. The Farnes' unifying theory does not allow the early existence of positive-mass dark matter, so dark matter needs modification.

To deal with greater than 50% of dark energy and the early existence of dark matter, this paper proposes the positive-negative mass universe pair model where the positive and the negative universes are symmetrical [8] [9] [10]. This paper starts with the proposed positive-negative-mass 11D membrane-antimembrane universe pair (without kinetic energy) which are transformed into the positive-negative-mass 10D string-antistring universe pair and the external dual gravities as in the Randall-Sundrum model [11] [12], resulting in the four equal and separate universes consisting of the positive-mass 10D universe, the positive-mass massive external gravity, the negative-mass 10D universe, and the negative-mass massive external gravity. The positive-mass 10D universe is transformed into 4D universe (home universe) with kinetic energy through the inflation and the Big Bang. Dark matter is the right-handed neutrino, exactly five times of baryonic matter in total mass. The other three universes without kinetic energy oscillate between 10D and 10D through 4D, resulting in the hidden universes when $D > 4$ and dark energy when $D = 4$ (the maximum dark energy = 3/4 = 75%). In this paper, the first modification is that 4D dark energy is created from the three oscillating dimensional universes (the positive-mass massive external gravity, the negative-mass universe, and the negative-mass massive external gravity) instead of only negative-mass universe in the Farnes' unifying theory. Otherwise, both the proposed dark energy and the Farnes' dark energy are the same.

To deal with the early existence of positive-mass dark matter in the CMB, the second modification is that the proposed dark matter in the CMB is same as the conventional positive-mass dark matter observed in the CMB. However, after the CMB, the proposed positive-mass dark matter turned into repulsive positive-mass dark matter when the condensed baryonic gas reached the critical surface density (derived from the acceleration constant a_0 in MOND [13] [14]) which induces dark matter repulsive force to transform dark matter in the region into repulsive dark matter repulsing one another, corresponding to the Farnes' repulsive negative-mass dark matter. Therefore, the proposed dark matter and the Farnes' dark matter have different beginnings, but both end up with the same repulsive force. In this way, a galaxy actually started with a dark matter core-cusp without repulsive dark matter. The removal of the core cusp by repulsive dark matter afterward solves many difficult problems in protogalaxy and galaxy evolutions.

The theoretical calculated percentages of dark energy, dark matter, and baryonic matter are 68.6 (as an input from the observation), 26, and 5.2, respectively, in agreement with observed 68.6, 26.5, and 4.9 [15], respectively. According to the calculation, dark energy started in 4.33 billion years ago in agreement

with the observed 4.71 ± 0.98 billion years ago [16]. In conclusion, the modified Farnes' unifying theory reinterprets the Farnes' equations, and is a unifying theory of dark energy, dark matter, and baryonic matter in the positive-negative mass universe pair. The unifying theory explains protogalaxy evolution and galaxy evolution in agreement with the observations.

In Section 2, unifying theory of dark energy and dark matter by Farnes is described. Section 3 explains the positive-negative mass universe pair. Section 4 describes protogalaxy and galaxy evolutions.

2. Unifying Theory of Dark Energy and Dark Matter

In the Farnes' unifying theory of dark energy and dark matter [1], dark energy and dark matter can be unified into a single negative-mass dark fluid which is created continuously through a creation tensor from the negative-mass universe to our positive-mass universe.

2.1. The Creation Tensor for Dark Energy and Dark Matter

In the Einstein's field equation,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{1}$$

where $R_{\mu\nu}$ is the Ricci curvature tensor, R is the scalar curvature, $g_{\mu\nu}$ is the metric tensor, Λ is the cosmological constant, G is Newton's gravitational constant, c is the speed of light in vacuum, and $T_{\mu\nu}$ is the stress-energy tensor.

In a homogeneous and isotopic universe, the Friedmann equation is

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2} \tag{2}$$

and the Friedmann acceleration equation.

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \tag{3}$$

where a is the scale factor, H is the Hubble parameter, ρ is the total mass density of the universe, p is the pressure, k is the curvature parameter or intrinsic curvature of space, and k/a^2 is the spatial curvature in any time-slice of the universe. k = +1, 0, and −1, indicate a closed, flat, and open universe respectively. In the positive mass universe corresponding to the standard matter-dominated universe solutions with a critical density given by $\rho_c = 3H^2/8\pi G$ and a total density parameter given by $\Omega = \rho/\rho_c$, where Ω = 1, <1, and >1, correspond to critical density, open, and closed universes respectively.

In the Farnes' unifying theory of dark energy and dark matter [1], baryonic matter has positive mass, while dark fluid consisting of dark energy and dark matter has negative masses. Dark fluid is created continuously through a "creation tensor" from the positive-mass dominated universe. Negative mass is created through the creation tensor, $C_{\mu\nu}$

$$T'_{\mu\upsilon} = T_{\mu\upsilon} + C_{\mu\upsilon} \tag{4}$$

Einstein's field equations are therefore modified to

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}\left(T_{\mu\upsilon} + C_{\mu\upsilon}\right) \tag{5}$$

The creation tensor generates negative mass density, ρ_-. Farnes proposed that the Friedmann equation can be written in terms of both the cosmological constant and negative mass density, ρ_- as follows.

$$\begin{aligned}\left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho_+ + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2} \\ &= \frac{8\pi G}{3}\rho_+ + \frac{8\pi G}{3}\rho_- - \frac{kc^2}{a^2}\end{aligned} \tag{6}$$

In this way, the cosmological constant (Λ) is equivalent to the negative-mass density, and $\Lambda = 8\pi G\rho_- / c^2$. The accelerated cosmic expansion is caused by negative mass which is created continuously through the creation tensor. As these negative masses can take the form of a cosmological constant, the field equations are modified to

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}\left(T^+_{\mu\upsilon} + T^-_{\mu\upsilon} + C_{\mu\upsilon}\right) \tag{7}$$

In the Farnes' unifying theory for dark matter, the positive-mass baryonic galaxy is surrounded by a halo of continuously-created negative masses, with constant density ρ_- and of total mass M_-. The positive mass particle is now immersed in a negative mass fluid that behaves with resemblance to a cosmological constant with $\Lambda = 8\pi G\rho_- / c^2$.

Negative-mass dark matter repulses each other, while dark matter still attracts to baryonic matter. The solution for the rotation curve with circular velocity, v, is deduced as follows.

$$\upsilon = \sqrt{\frac{GM_*}{r} - \frac{\Lambda c^2}{3} r^2} = \sqrt{\frac{GM_*}{r} - \frac{8\pi G\rho_-}{3} r^2} \tag{8}$$

where r is the distance from a central point mass $M_\star$ and ρ_- is the local negative mass density. For $\Lambda = 0$ or $\rho_- = 0$, the standard Keplerian rotation curve is obtained. However, for non-zero values of the cosmological constant or ρ_-, the rotation curve is modified. A negative cosmological constant flattens the rotation curve, causing a steady increase at larger galactic radii in agreement with the observed rotation curve.

In a N-body simulations of galaxy formation via negative-mass dark matter and positive-mass baryonic matter, Farnes starts with the galaxy that comprises 5000 particles, with a total positive mass of $M_+ = 1.0$. The initial positive mass particle distribution is located at the center of a cube of negative masses with volume 200^3. The initial conditions of these negative masses are set to be uniformly distributed in position and with zero initial velocity. The negative mass sea comprises 45,000 particles, with a total mass of $M_- = -3.0$. The negative masses at the outskirts of the cube are mutually-repelled by other surrounding negative masses and the cube begins to expand in volume. Meanwhile, the nega-

tive masses within the central portion of the cube are attracted towards the positive mass galaxy. From their initially zero velocities, the negative mass particles are slushed to-and-fro from either side of the positive mass galaxy. Eventually, the negative mass particles reach dynamic equilibrium in a halo that surrounds the positive mass galaxy and which extends out to several galactic radii. The negative mass particles have naturally formed a dark matter halo to have a flat central dark matter distribution with the constant dark matter density core.

With positive mass dark matter, the Navarro-Frenk-White (NFW) profile [2] derived from standard N-body simulations has the high density dark matter at the core of galaxy in disagreement with the observed distribution of dark matter. The high density dark matter at the core of galaxy is not observed [3]. This is known as the core-cusp problem or the cuspy halo problem and is currently unsolved [17]. A negative mass cosmology can therefore provide a solution to the cuspy halo problem by the repulsion within dark matter particles. This appears to be the only cosmological theory within the scientific literature that can explain and predict correctly the non-cuspy distribution of dark matter in galaxies from first principles of intrinsic repulsive dark matter. Dark matter of low-mass sterile neutrino and polytropic exponent under fermion degeneracy pressure also shows non-cuspy distribution [18].

2.2. The Modified Creation Tensor for Dark Energy and Dark Matter

This paper modifies and reinterprets the Farnes' unifying theory of dark energy and dark matter as negative-mass dark fluid derived from the positive-negative mass universe pair model where positive-mass and negative-mass universes are symmetrical. (The Farnes' unifying theory does not assume symmetrical positive-negative mass universes.) The first modification is about dark energy. With the symmetry between positive mass and negative mass, the portion of dark energy created through the creation tensor cannot start from zero to more than 50%. Since dark energy is 68.6%, dark energy is more than negative mass, and the creation tensor for dark energy needs modification. The first modification is to modify the origin of dark energy. The second modification is to modify dark matter. Evidence for early existence of dark matter comes from measurements on cosmological scales of anisotropies in the cosmic microwave background [5] [6] [7]. The CMB is the remnant radiation from the hot early days of the universe. The photons underwent oscillations that froze in just before decoupling from the baryonic matter at a redshift of 1100. The angular scale and height of the peaks (and troughs) of these oscillations are powerful probes of cosmological parameters, including the total energy density, the baryonic fraction, and the dark matter component. The CMB provides irrefutable evidence for the early existence of dark matter. The Farnes' unifying theory does not allow the early existence of positive dark matter, so the creation tensor for dark matter needs modification. The second modification is to modify the origin of dark matter.

3. The Positive-Negative Mass Universe Pair Model

To deal with more than 50% of dark energy and the early existence of dark matter in the CMB, this paper proposes as described previously [8] [9] [10] the positive-negative mass universe pair model where the positive and the negative universes are symmetrical. The positive-negative mass universe pair model involves the oscillating spacetime dimension number, the space structure, and the cyclic universe model.

3.1. The Oscillating Spacetime Dimension Number

The oscillating spacetime dimension number oscillates between 11D membrane and 11D membrane through 10D string and between 10D particle and 10D particle through 4D particle as described in the previous papers [19] [20]. Because membrane and string require 11D and 10D, respectively, any spacetime dimension below 10D has to be the spacetime dimension for particle which allows any spacetime dimension numbers.

3.1.1. The Membrane-String Oscillation between 11D4d and 11D4d

In the membrane-string oscillation, the 11D brane is transformed into the 10D string with an extra dimension. This 11D warped brane world in the 10D string with compact extra dimension is analogous to the 5D warped brane world in our 4D universe with compact extra dimension in the Randall-Sundrum model [11] [12]. The RS1 of the Randall-Sundrum model produces the two different branes consisting of the Tevbrane and the Planckbrane. The Planckbrane has very strong gravity, while the Tevbrane has all other forces and extremely weak gravity. There are boundaries between the Tevbrane and the Planckbrane. The Tevbrane corresponds to our universe with extremely weak gravity comparing with other forces. Almost all gravity is external gravity outside of our universe. The Randall-Sundrum model explains the hierarchy problem between gravity and other forces.

In the same way as the RS1 in the Randall-Sundrum model, two 11D membranes produce the 10D string corresponding to the Tevbrane and the external gravity corresponding to the Planckbrane. The 11D membrane, the 10D string, and the external gravity have about the same energy. The 10D string has the extremely weak internal gravity as in the Tevbrane with an extremely weak gravity in the Randall-Sumdrum model.

The RS1 in the Randall-Sundrum model

$$\begin{aligned} &\text{5D brane world} \xrightarrow{\text{compact extra space dimension}} \text{Tevbrane} + \text{Planckbrane} \\ &\text{two 11D membranes} \xleftarrow{\text{oscillation between 11D membrane and 10D string}} \text{10D string} + \text{external gravity} \end{aligned} \tag{9}$$

The membrane-string oscillation is reversible, so the 10D string and the external gravity can also reverse back to the 11D membrane.

3.1.2. The Particle Oscillation between 10D and 10D through 4D

As described previously [8] [9] [10], the space-time dimension numbers oscillate

reversibly between 10D and 4D reversibly dimension by dimension without compactization. The oscillating space-time numbers from 10D to 4D relate to varying speed of light. Varying speed of light has been proposed to explain the horizon problem of cosmology [21] [22]. J. D. Barrow [23] proposes that the time dependent speed of light varies as some power of the expansion scale factor a in such way that

$$c(t) = c_0 a^n, \tag{10}$$

where c is the speed of light and n are constants. The increase of speed of light is continuous.

In this paper, the speed of light is invariant in a constant space-time dimension number, and the speed of light varies with varying space-time dimension number from 4 to 10 as follows.

$$c_{\mathrm{D}} = c/\alpha^{\mathrm{D}-4}, \tag{11}$$

where c is the observed speed of light in the 4D space-time, c_{D} is the quantized varying speed of light in space-time dimension number, D, from 4 to 10, and α is the fine structure constant for electromagnetism. The speed of light increases with the increasing space-time dimension number D. Since the speed of light for >4D particle is greater than the speed of light for 4D particle, the observation of >4D particles by 4D particles violates casualty. Thus, >4D particles are hidden particles with respect to 4D particles. Particles with different space-time dimensions are transparent and oblivious to one another, and separate from one another if possible.

As described previously [8] [9] [10], the particle oscillation between 10D and 10D through 4D involves mass dimension (denoted as d) to represent the mass. In the initial condition for the oscillation, D + d = 14 where D and d are between 4 and 10. For an example, a dimension has a dual spacetime-mass dimension numbers of 10D4d or 4D10d. The transformations for oscillating dimension number between 10D and 4D consist of the varying speed of light dimensional (VSLD) transformation for spacetime dimension D and the varying supersymmetry dimensional (VSD) transformation for mass dimension d. For the VSLD transformation for D, $E = M_0c^2$ modified by Equation (12) is expressed as

$$E = M_0 c_{\mathrm{D}}^2 = M_0 \cdot \left(c^2/\alpha^{2(\mathrm{D}-4)}\right), \tag{12}$$

$$M_{0,\mathrm{D},\mathrm{d}} = M_{0,\mathrm{D}-n,\mathrm{d}+n}\alpha^{2n}, \tag{13}$$

$$E_{\mathrm{vacuum,D}} = E - M_{0,\mathrm{D}}c^2, \tag{14}$$

$$\mathrm{D},\mathrm{d} \xrightarrow{\text{VSLD transformation}} (\mathrm{D} \mp n), (\mathrm{d} \pm n) \tag{15}$$

where c_{D} is the quantized varying speed of light in space-time dimension number, D, from 4 to 10, c is the observed speed of light in the 4D space-time, α is the fine structure constant for electromagnetism, E is energy, M_0 is rest mass, D is the space-time dimension number from 4 to 10, d is the mass dimension number from 4 to 10, n is an integer, and E_{vacuum} = vacuum energy. From Equa-

tion (12), 10D has the lowest rest mass, and 4D has the highest rest mass. According to the calculation from Equation (13), the rest mass of 4D is $1/\alpha^{12} \approx 137^{12}$ times of the mass of 10D. From Equation (14), 10D has the highest vacuum energy, while 4D particle has zero vacuum energy. A particle with 10D is transformed to a particle with 4D from Equation (15) through the VSLD transformation. Spacetime dimension number decreases with decreasing speed of light, decreasing vacuum energy, and increasing rest mass. The 4D and the 10D have zero and the highest vacuum energies, respectively.

In the normal supersymmetry transformation, the repeated application of the fermion-boson supersymmetry transformation carries over a boson (or fermion) from one point to the same boson (or fermion) at another point at the same mass, resulting in translation without changing mass. Under the varying supersymmetry dimensional (VSD) transformation, the repeated application of the fermion-boson supersymmetry transformation carries over a boson from one point to the boson at another point at different mass dimension number at different mass, resulting in translation and fractionalization or condensation. The repeated VSD transformation carries over a boson B_d into a fermion F_d and a fermion F_d to a boson B_{d-1}, which can be expressed as follows.

$$M_{d,F} = M_{d,B}\alpha_{d,B}, \tag{16}$$

$$M_{d-1,B} = M_{d,F}\alpha_{d,F}, \tag{17}$$

where $M_{d,B}$ and $M_{d,F}$ are the masses for a boson and a fermion, respectively, d is the mass dimension number, and $\alpha_{d,B}$ or $\alpha_{d,F}$ is the fine structure constant that is the ratio between the masses of a boson and its fermionic partner. where $M_{d,B}$ and $M_{d,F}$ are the masses for a boson and a fermion, respectively, d is the mass dimension number, and $\alpha_{d,B}$ or $\alpha_{d,F}$ is the fine structure constant that is the ratio between the masses of a boson and its fermionic partner. Assuming α's are the same, it can be expressed as

$$M_{d,B} = M_{d+1,B}\alpha_{d+1}^2. \tag{18}$$

The oscillating dimension number transformation between 10D4d and 10D4d through 4D4d involves both the VSLD transformation and the VSD transformation as the stepwise two-step transformation as follows.

stepwise two-step varying transformation

$$\begin{aligned} &(1)\quad \mathrm{D,d} \xleftrightarrow{\text{VSLD}} (\mathrm{D}\mp 1), (\mathrm{d}\pm 1) \\ &(2)\quad \mathrm{D,d} \xleftrightarrow{\text{VSD}} \mathrm{D}, (\mathrm{d}\pm 1) \end{aligned} \tag{19}$$

The repetitive stepwise two-step dimension number oscillation between 10D4d and 10D4d through 4D4d as follows.

$$\begin{aligned} &10\mathrm{D4d} \to 9\mathrm{D5d} \to 9\mathrm{D4d} \to 8\mathrm{D5d} \to 8\mathrm{D4d} \to 7\mathrm{D5d} \to 7\mathrm{D4d} \to 6\mathrm{D5d} \\ &\to 6\mathrm{D4d} \to 5\mathrm{D5d} \to 5\mathrm{D4d} \to 4\mathrm{D5d} \to 4\mathrm{D4d} \to 5\mathrm{D4d} \to 5\mathrm{D5d} \to 6\mathrm{D4d} \\ &\to 6\mathrm{D5d} \to 7\mathrm{D4d} \to 7\mathrm{D5d} \to 8\mathrm{D4d} \to 8\mathrm{D5d} \to 9\mathrm{D4d} \to 9\mathrm{D5d} \to 10\mathrm{D4d} \end{aligned} \tag{20}$$

From Equation (18), the mass of 9D4d is $\alpha^2 \approx (1/137)^2$ times of the mass of 9D5d through the varying supersymmetry transformation. The transformation

from a higher mass dimensional particle to the adjacent lower mass dimensional particle is the fractionalization of the higher dimensional particle to the many lower dimensional particles in such way that the number of lower dimensional particles becomes

$$N_{\mathrm{d}-1} = N_{\mathrm{d}} / \alpha^2 \approx N_{\mathrm{d}} (137)^2 \tag{21}$$

The fractionalization also applies to D for 10D4d to 9D4d, so

$$N_{\mathrm{D}-1} = N_{\mathrm{D}} / \alpha^2 \tag{22}$$

Since the supersymmetry transformation involves translation, this stepwise varying supersymmetry transformation leads to a translational fractionalization, resulting in the cosmic expansion. Afterward, the QVSL transformation transforms 9D4d into 8D5d with a higher mass. The two-step transformation repeats until 4D4d, and then reverses stepwise back to 10D4d for the cosmic contraction. The oscillation between 10D and 4D results in the reversible cyclic fractionalization-contraction for the reversible cyclic expansion-contraction of the universe which does not involve irreversible kinetic energy.

3.2. The Space Structure

In the space structure, attachment space that attaches matter to the space relates to rest mass, and detachment space that detaches matter from the space relates to kinetic energy. Attachment space is the space precursor of the transitional Higgs field, and detachment space is the space precursor of the transitional reverse Higgs field.

3.2.1. The Higgs Mechanism

In conventional physics, space does not couple with particles, and is the passive zero-energy ground state space. Under spontaneous symmetry breaking in conventional physics, the passive zero-energy ground state is converted into the active, permanent, and ubiquitous nonzero-energy Higgs field, which couples with massless particle to produce the transitional Higgs field-particle composite. Under spontaneous symmetry restoring, the transitional Higgs field-particle composite is converted into the massive particle with the longitudinal component on zero-energy ground state without the Higgs field as follows.

$$\begin{aligned}
&\text{zero-energy groud state space} \xrightarrow{\text{spontaneous symmetry breaking}} \\
&\text{nonzero-energy scalar Higgs field} \xrightarrow{\text{massless particle}} \\
&\left[\text{the transitional nonzero-energy Higgs field-particle composite}\right] \\
&\xrightarrow{\text{spontaneous symmetry restoring}} \text{massive particle with the longitudinal component} \\
&\text{on zero-energy ground state space without the Higgs field}
\end{aligned} \tag{23}$$

In conventional physics, the nonzero-energy scalar Higgs Field exists permanently in the universe. The problem with such nonzero-energy field is the cosmological constant problem from the huge gravitational effect by the nonzero-energy Higgs field in contrast to the observation [24].

Unlike passive space in conventional physics, attachment space actively couples to massless particle. Under spontaneous symmetry breaking, attachment space as the active zero-energy ground state space couples with massless particle to form momentarily the transitional non-zero energy Higgs field-particle composite. The Higgs field is momentary and transitional, avoiding the cosmological constant problem. Under spontaneous symmetry restoring, the transitional nonzero-energy Higgs field-particle composite is converted into massive particle with the longitudinal component on zero-energy attachment space without the Higgs field as follows.

$$\begin{aligned}&\text{massless particle} + \text{zero-energy attachment space} \xrightarrow{\text{spontaneous symmetry breaking}} \\ &\left[\text{the transitional non-zero energy Higgs field-particle composite}\right] \\ &\xrightarrow{\text{spontaneous symmetry restoring}} \text{massive particle with the longitudinal component} \\ &\text{on zero-energy attachment space without the Higgs field}\end{aligned} \tag{24}$$

Detachment space is the space precursor of the reverse Higgs field. Unlike the conventional model, detachment space actively couples to massive particle. Under spontaneous symmetry breaking, the coupling of massive particle to zero-energy detachment space produces the transitional nonzero-energy reverse Higgs field-particle composite which under spontaneous symmetry restoring produces massless particle on zero-energy detachment space without the longitudinal component without the reverse Higgs field as follows.

$$\begin{aligned}&\text{massive particle} + \text{zero-energy detachment space} \xrightarrow{\text{spontaneous symmetry breaking}} \\ &\left[\text{the transitional nonzero-energy reverse Higgs field-particle composite}\right] \\ &\xrightarrow{\text{spontaneous symmetry restoring}} \text{massless particle without the longitudinal component} \\ &\text{on zero-energy detachment space without the reverse Higgs field}\end{aligned} \tag{25}$$

For the electroweak interaction in the Standard model where the electromagnetic interaction and the weak interaction are combined into one symmetry group, under spontaneous symmetry breaking, the coupling of the massless weak W, weak Z, and electromagnetic A (photon) bosons to zero-energy attachment space produces the transitional nonzero-energy Higgs fields-bosons composites which under partial spontaneous symmetry restoring produce massive W and Z bosons on zero-energy attachment space with the longitudinal component without the Higgs field, massless A (photon), and massive Higgs boson as follows.

$$\begin{aligned}&\text{massless WZ} + \text{zero-energy WZ attachment space} + \text{massless A} \\ &+ \text{zero-energy A attachment space A} \xrightarrow{\text{spontaneous symmetry breaking}} \\ &\left[\text{the transitional nonzero-energy WZ Higgs field} - \text{WZ composite}\right] \\ &+ \left[\text{nonzero-energy A Higgs field} - \text{A composite}\right] \xrightarrow{\text{partial spontaneous symmetry restoring}} \\ &\text{massive WZ with the longitudinal component on attachment space without} \\ &\text{the Higgs field} + \text{massless A} + \text{the nonzero energy massive Higgs boson}\end{aligned} \tag{26}$$

In terms of mathematical expression, the conventional permanent Higgs field model and the transitional Higgs field model are identical. The interpretations of

the mathematical expression are different for the permanent Higgs field model and the transitional Higgs field model. The transitional Higgs field model avoids the cosmological problem in the permanent Higgs field model.

In the Higgs mechanism, gauge bosons are assumed to be massless originally. Elementary fermions (leptons and quarks) can be assumed to be massive originally. However, the observed neutrinos are nearly massless and left-handed only. The paper posits that the left handed became massless through the reverse Higgs mechanism. For the symmetrical massive left handed neutrinos and right-handed neutrinos under spontaneous symmetry breaking, the coupling of the massive left handed neutrinos and the massive right handed neutrinos to zero-energy detachment space produces the transitional nonzero-energy reverse Higgs fields-neutrinos composites which under partial spontaneous symmetry restoring produce massless left handed neutrinos on zero-energy detachment space without the longitudinal component without the reverse Higgs field, massive right-handed neutrinos (dark matter), and the massive reverse Higgs boson as follows.

$$\begin{aligned}&\text{massive } \nu_L + \text{zero energy } \nu_L \text{ detachment space} + \text{massive } \nu_R\\&+ \text{zero-energy } \nu_R \text{ detachment space} \xrightarrow{\text{spontaneous symmetry breaking}}\\&\left[\text{the transitional nonzero-energy } \nu_L \text{ reverse Higgs field} - \nu_L \text{ composite}\right]\\&+\left[\text{nonzero-energy } \nu_R \text{ reverse Higgs field} - \nu_R \text{ composite}\right]\\&\xrightarrow{\text{partial spontaneous symmetry restoring}} \text{massless } \nu_L \text{ without the longitudinal}\\&\text{component on detachment space without the Higgs field}\\&+\text{massive } \nu_R + \text{the nonzero energy massive reverse Higgs boson}\end{aligned} \tag{27}$$

As described in the previous paper [25], the reverse Higgs boson was observed as the two unusual steeply upward-going ultra-high-energy (UHE) cosmic ray events with energies of ≈0.6 EeV [26] and ≈0.56 EeV [27] in the Antarctic Impulsive Transient Antenna (ANITA) experiment [28]. These shower events have the characteristics of the decay of a tau lepton, which emerges from the surface of the ice, and the tau lepton is explained as the product of a UHE parent tau neutrino by the charged-current interactions with the Earth matter. However, such UHE tau neutrino cannot survive the passage through the Earth. The previous paper posits that the upward-going ANITA events are derived from the cosmic ray of the baryonic-dark matter reverse Higgs boson that travels through the Earth. The calculated value for the reverse Higgs boson is 0.47 EeV in good agreement with the observed 0.56 and 0.6 EeV. As shown in Section 3.3., dark matter (sterile neutrinos) is part of the periodic table of elementary particles for baryonic matter and dark matter.

3.2.2. The Three Spaces

The combination of n units of attachment space as 1 and n units of detachment space as 0 brings about three different spaces: binary partition space, miscible space, or binary lattice space as below.

$$\underset{\text{attachment space}}{(1)_n} + \underset{\text{detachment space}}{(0)_n} \xrightarrow{\text{combination}} \underset{\text{binary partition space}}{(1)_n(0)_n}, \underset{\text{miscible space}}{(1+0)_n} \text{ or } \underset{\text{binary lattice space}}{(1\,0)_n} \tag{28}$$

Binary partition space, $(1)_n(0)_n$, consists of two separated continuous phases of multiple quantized units of attachment space and detachment space, and it is the space structure for wave-particle duality in quantum mechanics. In miscible space, $(1+0)_n$, attachment space is miscible to detachment space, and there is no separation of attachment space and detachment space, and it is the space structure for miscible mass-energy in relativity. Binary lattice space, $(1\ 0)_n$, consists of repetitive units of alternative attachment space and detachment space, and it is the space structure for virtual particles in quantum field theory.

An object in binary partition space $(1)_n(0)_n$ has both movement and rest at the same time, resulting in wave-particle duality for movement-rest duality in quantum mechanics. An object in binary partition space cannot be completely at movement (zero momentum) or completely at rest (zero distance), resulting in the uncertainty principle as follows.

$$\sigma_x \sigma_p \geq \frac{\hbar}{2} \tag{29}$$

where x is position and p is momentum. The interference to binary partition space collapses binary partition space, resulting in miscible space as follows.

$$\underset{\text{binary partition space}}{(0)_n(1)_n} \xrightarrow{\text{collapse}} \underset{\text{miscible space}}{(0+1)_n} \tag{30}$$

In miscible space, attachment space is miscible to detachment space, resulting in miscible mass and energy where attachment space for mass provides zero speed for rest mass m_0, while detachment space for energy provides the speed of light for kinetic energy. The total energy is the combination of both as follows.

$$E = K + m_0c^2 = \gamma m_0c^2 \tag{31}$$

where $\gamma = 1/\left(1-v^2/c^2\right)^{1/2}$ is the Lorentz factor for time dilation, m_0 is rest mass, E is the total energy, and K is the kinetic energy. Binary lattice space, $(1\ 0)_n$ is the space structure for virtual particles in quantum field theory which will be described in the next section.

3.3. The Cyclic Universes Model

The seven steps in the cyclic universes model are 1) the formation of positive-mass and negative-mass dual 11D4d membrane-antimembrane universes from the zero-mass inter-universal void, 2) the transformation of the 11D4d membrane-antimembrane dual universes to the 10D4d string-antistring dual universes and dual external dual gravities, 3) the transformation from the string-antistring dual universes to the particle-antiparticle dual universes, 4) the transformation of the positive-mass 10D4d universe into the positive-mass 4D universe, and the transformation of the other three universes into the hidden oscillating dimension number universes from 10D to 5D, 5) the transformation of all four universes into the 4D universes, 6) the positive-mass 4D universe and the three hidden oscillating dimension number from 5D to 10D, and 7) the return to the 10D4d particle-antiparticle universes (the step 3) as in **Figure 1** and **Figure 2**.

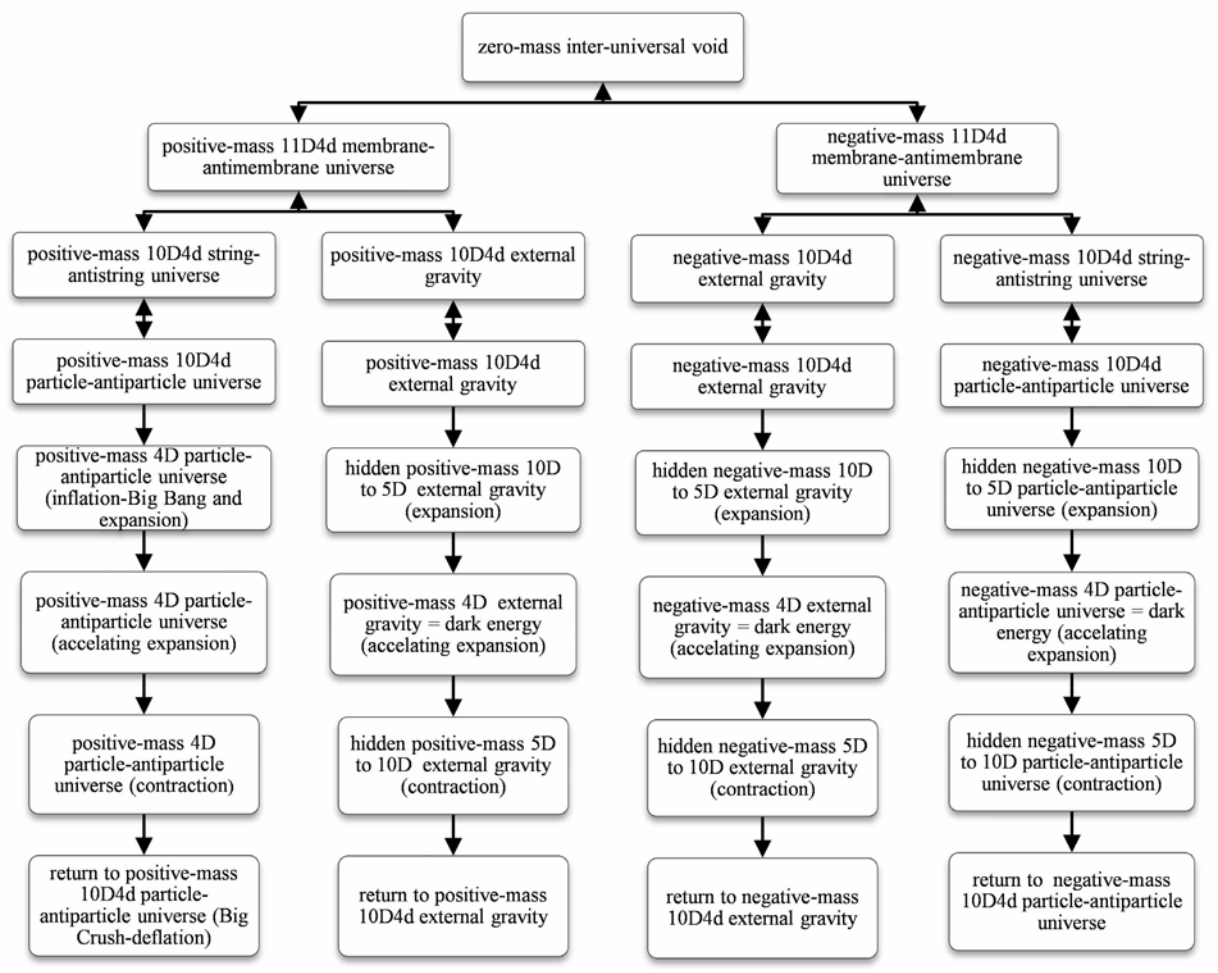


Figure 1. The cyclic universes model.

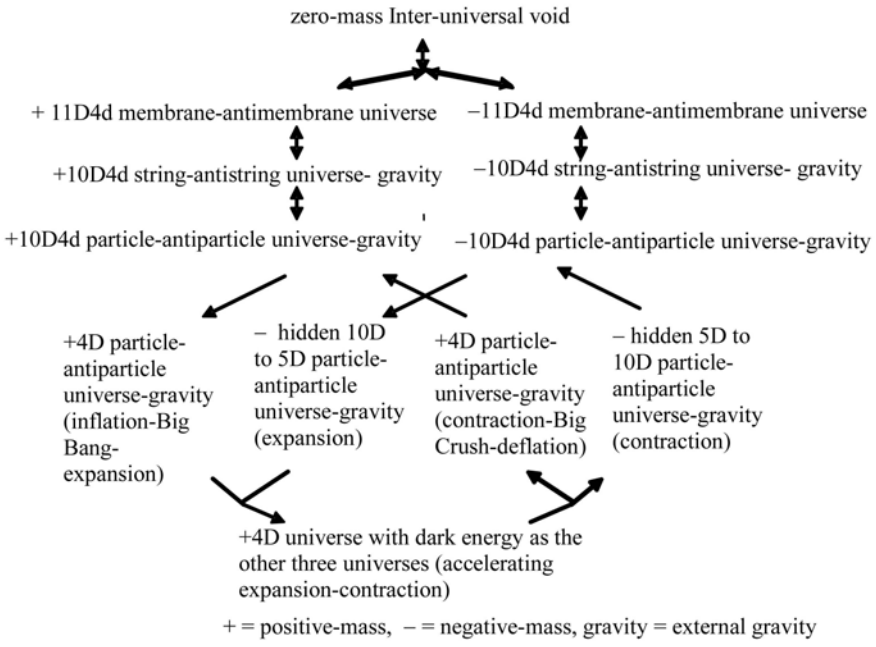


Figure 2. The cyclic universes model.

1) the formation of the positive-negative-mass 11D4d membrane-antimembrane dual universes

In the cyclic universes model, the universes start with the positive-mass 11D4d membrane-antimembrane universe and the negative-mass 11D4d membrane-antimembrane universe derived from the zero-mass inter-universal void. The mass sum of the dual universes is zero. The zero-mass inter-universal void and the dual universe are reversible, so the dual universe can reverse back to the zero-mass inter-universal void. The inter-universal void contains only detachment space to prevent irreversible inter-universal collision, while the dual universes contain only attachment space without kinetic energy. These 11D dual universes are the universes with the oscillating spacetime dimension number, and start the process of oscillation between 11D and 11D through 10D and between 10D and 10D through 4D.

2) the transformation of the 11D4d membrane-antimembrane dual universes to the 10D4d string-antistring dual universes and dual external gravities

As described in Section 3.1.1., the transformation of the 11D membrane produces the 10D string and the external gravity. The results are the positive-mass 10D4d string-antistring universe, the positive-mass external gravity, the negative-mass 10D4d string-antistring universe, and the negative-mass external gravity. These four universes are separate, and have equal energy. The 10D4d string-external gravity and the 11D4d membrane are reversible.

3) the transformation from the string-antistring dual universes to the particle-antiparticle dual universes

Since string exists only in 10D, so any further transformation of D to lower than 10 cannot be string. As a result, to transform lower than 10, string-antistring is converted into particle-antiparticle. The results are the positive-mass 10D4d particle-antiparticle universe, the positive-mass 10D4d external gravity, the negative-mass 10D4d particle-antiparticle universe, and the negative-mass 10D4d external gravity. Particle-antiparticle and string-antistring are reversible.

4) the transformation of the positive-mass 10D4d particle-antiparticle universe into the positive-mass 4D universe, and the transformation of other three universes into the hidden oscillating dimension number universes from 10D to 5D

The positive-mass 10D particle-antiparticle universe is transformed into the positive-mass 4D universe to produce the 4D standard model particles. Under the oscillating spacetime dimension number, the negative-mass 10D particle-antiparticle, the positive-mass external gravity, and the negative-mass external gravity are transformed into the hidden oscillating dimension number universes from 10D to 5D.

4a) the formation of the positive-mass 4D particle-antiparticle universe

The formation of the positive-mass 4D particle-antiparticle universe includes the inflation and followed by the Big Bang. The inflation involves the VSLD transformation from 10D4d to 4D10d, because from Equation (13), the rest mass

M_0 of 4D10d is $M_{0,10} = M_{0,4} / \alpha^{2(10-4)} \approx 137^{12}$ times of the rest mass of 10D4d, resulting in the inflation for the rapid expansion

The Big Bang involves the entrance of detachment space from the inter-universal void to the positive-mass universe. Detachment space introduces massless particles and kinetic energy for the cosmic expansion, and forms the three spaces with attachment space. The Big Bang consists of the two steps. In the first step, all particles are converted into massless particles by detachment space for the reverse Higgs field as in Equation (25). The second step involves the partial conversion of massless particles into massive particles by attachment space for the Higgs field to produce massless particles, massive particles, and the Higgs boson as in Equations (26) and (27) for the standard model. The emergence of detachment space starts kinetic energy which causes the cosmic expansion as the Big Bang.

$$\begin{gathered}\text{the inflation and the Big Bang}\left(\text{detachment space}+\text{partial attachment space}\right)\\ 10\text{D4d}\xrightarrow{\text{the inflation}}4\text{D10d}\xrightarrow{\text{detachment space(reverse Higgs field)}}\text{massless particles}\\ \xrightarrow{\text{partial attachment space(Higgs field)}}\text{massive particles, massless particles, Higgs boson}\end{gathered} \tag{32}$$

4D10d particle was sliced into six different particles: 4D9d, 4D8d, 4D7d, 4D6d, 4D5d, and 4D4d equally by mass. Baryonic matter is 4D4d, while dark matter consisted of the other five types of particles (4D9d, 4D8d, 4D7d, 4D6d, and 4D5d) as follows.

$$\begin{gathered}10\text{D4d}\xrightarrow{\text{cosmic inflation}}4\text{D10d}\xrightarrow{\text{the slicing}}\\ \text{baryonic matter}\left(4\text{D4d}\right)+\text{dark matter}\left(4\text{D5d},4\text{D6d},4\text{D7d},4\text{D8d},4\text{D9d}\right)\end{gathered} \tag{33}$$

As a result, the mass ratio of dark matter to baryonic matter is 5 to 1. At observed 68.6% dark energy [15], the calculated values for baryonic matter and dark matter (with the 1:5 ratio) are 5.2% (=(100 - 68.6)/6) and 26 % (=5.2 × 5), respectively, in good agreement with observed 4.9% and 26.5%, respectively [15].

As described in Section 3.2.2., the space structure with both attachment space and detachment produces binary partition space $(1)_n(0)_n$ for wave-particle duality in quantum mechanics, miscible space $(1+0)_n$ for miscible mass-energy in relativity, and binary lattice space $(1\ 0)_n$ for virtual particles in quantum field theory. Binary lattice space is derived from the slicing of mass dimensions by detachment space as described by Bounias and Krasnoholovets [29] who propose another explanation of the reduction of >4 D space-time into 4D space-time by slicing > 4D space-time into infinitely many 4D quantized units surrounding the 4D core particle. Such slicing of >4D space-time is like slicing 3-space D object into 2-space D object in the way stated by Michel Bounias as follows: "You cannot put a pot into a sheet without changing the shape of the 2-D sheet into a 3-D dimensional packet. Only a 2-D slice of the pot could be a part of sheet". 4D10d particles emerge after the inflation. 10d mass dimension is sliced by detachment space into 9d, 8d, 7d, 6d, 5d, and 4d plus the binary lattice space $(1\ 0)_n$ for virtual particles in quantum field theory. For an example, the slicing of 10d particle into 4d particle is as follows.

$$1_{10} \xrightarrow{\text{slicing}} 1_4 \qquad \sum_{d=5}^{10}\left(0_4\ 1_4\right)_{n,d} \tag{34}$$

10d particle 4d core particle binary lattice space

where 1 is attachment space, 0 is detachment space, 1_{10} is 10d particle, 1_4 is 4d particle, d is the mass dimension number of the dimension to be sliced, n as the number of slices for each dimension, and $(0_4\ 1_4)_n$ is binary lattice space with repetitive units of alternative 4d attachment space and 4d detachment space. For 4d particle starting from 10d particle, the mass dimension number of the dimension to be sliced is from d = 5 to d = 10. Each mass dimension is sliced into infinite quantized units ($n = \infty$) of binary lattice space, $(0_4\ 1_4)_\infty$. For 4d particle, the 4d core particle is surrounded by 6 types (from d = 5 to d = 10) of infinite quantized units of binary lattice space. Such infinite quantized units of binary lattice space represent the infinite units ($n = \infty$) of separate virtual orbitals for virtual particles in a gauge force field, while the dimension to be sliced is "mass dimensional orbital" (DO), representing a type of gauge force field. In addition to the six DO's for gauge force fields from d = 5 to d = 10, the weak internal gravity appears as the seventh DO at d = 11. As a result, there are seven mass dimensional orbitals as in **Figure 3**.

The seven mass dimensions are arranged as $F_5\ B_5\ F_6\ B_6\ F_7\ B_7\ F_8\ B_8\ F_9\ B_9\ F_{10}\ B_{10}\ F_{11}\ B_{11}$, where F_d and B_d are mass dimensional fermion and mass dimensional boson, respectively. Under the varying supersymmetry dimensional (VSD) transformation, the mass of mass dimensional fermion and the mass of mass dimensional boson are related to each other with three simple formulas as the follows.

$$M_{d,B} = M_{d,F}/\alpha_d \tag{35}$$

$$M_{d+1,F} = M_{d,B}/\alpha_{d+1} \tag{36}$$

$$M_{d+1,B} = M_{d,B}/\alpha_{d+1}^2, \tag{37}$$

where d is the mass dimension number, F is fermion, and B is boson. Each dimension has its own α_d, and all α_d's except α_7 (α_w) of the seventh dimension (weak interaction) are equal to α, the fine structure constant of electromagnetism.

As shown in the previous paper "Split Membrane 11D Spacetime = 1D Eleventh Dimension Interval Space + 6D Rishon Space + 3D Higgs Space + 1D Einstein Time: Cosmology" [10], quarks are derived from the split 11D membrane. To represent quarks, the additional seven mass dimensions as the "auxiliary mass dimensions" (a's) to represent quarks and unstable leptons in addition to the seven "principal mass dimensions" (d's) to represent neutrino-electron and gauge bosons as in **Figure 4**.

The Periodic Table of Elementary Particles

The periodic table of elementary particles [25] [30] is based on the seven principal mass dimensions (d's) for stable baryonic matter leptons (electron and neutrinos), gauge bosons (all forces), gravity, and dark matter (five sterile dark

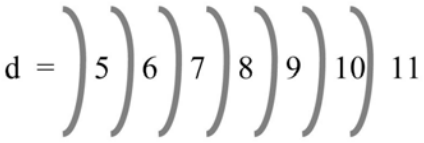


Figure 3. The seven mass dimensions as mass dimensional orbitals.

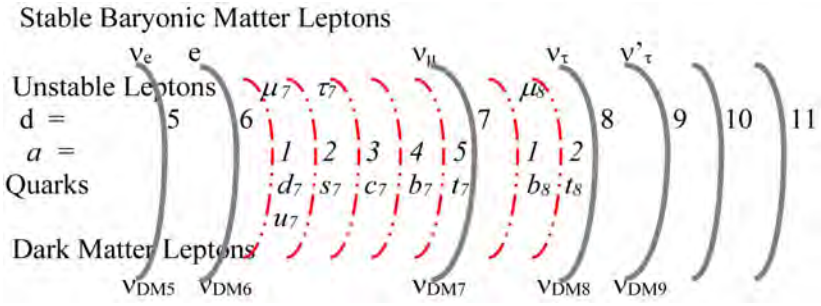


Figure 4. The seven principal mass dimensions (solid lines) denoted by the principal mass dimension number d and the seven auxiliary mass dimensions (dash-dotted lines) denoted by the auxiliary mass dimension number *a*.

matter neutrinos) and the seven auxiliary mass dimensions (a's) for unstable leptons (muon and tau) and quarks (d, u, s, c, b, and t) as in **Figure 4** and **Table 1**.

In the periodic table of elementary particles, the five dark matter particles are derived from Equation (33). Without electromagnetism at d = 5, dark matter does not have charge particle, and has to be neutrinos. Initially derived from Equation (33) and the symmetry between dark matter and baryonic matter, there were five dark matter massive right-handed neutrinos and one baryonic matter massive left-handed neutrino. Through the reverse Higgs mechanism as Equation (27), the left-handed neutrino becomes massless, while the right-handed neutrinos as sterile dark matter neutrinos remain massive. The reverse Higgs boson was observed [24].

As sterile neutrinos, dark matter does not react with baryonic matter possibly except baryonic neutrinos. A new observation of Excess Electronic Recoil Events in XENON1T indicates an excess over known backgrounds is observed below 7 keV, rising towards lower energies and prominent between 2 - 3 keV which enables the most sensitive searches for solar axions, an enhanced neutrino magnetic moment using solar neutrinos, and bosonic dark matter [31]. This event can be the rare reaction between massive ν_{DM} and massless ν_{BM} to produce boson B_5 at 3.7 keV (**Table 2**).

$$\text{massive } \nu_{\mathrm{DM}} + \text{massless } \bar{\nu}_{\mathrm{BM}} \rightarrow \text{boson } \mathrm{B}_5 \text{ at 3.7 keV} \tag{38}$$

All neutrinos and electron as well as gauge bosons are in the principal mass dimensions. All quarks and unstable leptons are in the auxiliary mass dimensions. The three generations of baryonic matter lepton-quark is the maximum generations allowed for the seven principal dimensions and the seven auxiliary dimensions.

Table 1. The periodic table of elementary particles for baryonic matter and dark matter. d = principal mass dimension number, a = auxiliary mass dimension number, DM = dark matter, BM = baryonic matter.

d	a = 0	a = 0	1	2	1	2	3	4	5	a = 0
	Stable Baryonic Matter Leptons	Dark Matter Leptons	Unstable Leptons		Quarks					Bosons
5	ν_e	ν_{DM5}								B_5 = A electromagnetism
6	e	ν_{DM6}								B_6 = g* strong (basic gluon for quarks)
7	ν_μ	ν_{DM7}	μ_7	τ_7	d_7/u_7	s_7	c_7	b_7	t_7	$B_7 = Z^0_L$ left-handed BM weak
8	ν_τ	ν_{DM8}	μ_8 (absent)		b_8 (absent)	t_8				$B_8 = Z^0_R$ right-handed DM weak
9	ν'_τ (high-mass ν_τ)	ν_{DM9}								B_9 = dark matter repulsive force
10										B_{10} = particle-antiparticle asymmetry
11	gravitino									B_{11} = gravity

Table 2. The masses of the principal mass dimensions (gauge bosons).

B_d	M_d	GeV (calculated)	Gauge boson	Interaction
B_5	$M_e\alpha$	3.7×10^{-6}	A = photon	Electromagnetic
B_6	M_e/α	7×10^{-2} (70.02 MeV)	g* = basic gluon	Strong
B_7	$M_Z = M_{B6}/\alpha_w^2$	91.1876 (given)	Z_L	weak (left) for baryonic matter
B_8	$M_7/\alpha^2 = M_Z/\alpha^2$	1.71×10^6	Z_R	weak (right) for dark matter
B_9	$M_8/\alpha^2 = M_Z/\alpha^4$	3.22×10^{10}		dark matter repulsive force
B_{10}	$M_9/\alpha^2 = M_Z/\alpha^6$	6.04×10^{14}		particle-antiparticle asymmetry
B_{11}	$M_{10}/\alpha^2 = M_Z/\alpha^8$	1.13×10^{19}	G	gravity

Gauge Bosons

In the periodic table of elementary particles, the given observed masses are the mass of electron for F_6 and the mass of Z boson for B_7. From Equations (35), (36) and (37), $\alpha_w = \alpha_7 = \alpha$ of week interaction $= (M_{B6}/M_{B7})^{1/2} = (M_{F6}/\alpha/M_{B7})^{1/2} = (M_e/\alpha/M_Z)^{1/2} = 0.02771$. Therefore, the masses of gauge bosons are as in Table 2.

The lowest energy gauge boson (B_5) at d = 5 is the Coulomb field for electromagnetism. The second gauge lowest boson (B_6) at d = 6 is basic gluon (g* = 70 MeV ≈ one half of pion) is the strong force as the nuclear force in the pion theory [32] where pions mediate the strong interaction at long enough distances (longer than the nucleon radius) or low enough energies. B_6 is denoted as basic gluon, g*. At short enough distances (shorter than the nucleon radius) or high enough energies, gluons emerge to confine fractional charge quarks. Fractional charge quarks are confined by gluons in QCD (quantum chromodynamics). No

isolated fractional charge quark is allowed, and only collective integer charge quark composites are allowed. In general, collective fractional charges are confined by the short-distance confinement force field where the sum of the collective fractional charges is integer [33]. As a result, fractional charges are confined and collective. The confinement force field includes gluons for collective fractional charge quarks in hadrons and the magnetic flux quanta for collective fractional charge quasiparticles in the fractional quantum Hall effect (FQHE) [34] [35] [36].

The third lowest boson (B_7) at d = 7 is Z_L for the left-handed weak interaction among leptons and quarks. Spontaneous symmetry breaking produces massive weak bosons, massless photon and the Higgs boson as Equation (27). Massive weak bosons produce short-distance interaction. B_8 at d = 8 is Z_R for the right-handed weak interaction among dark matter neutrinos as dark matter neutrino oscillation. The symmetry between Z_R and Z_L provides the neutrino oscillation for both baryonic matter neutrinos [37] and dark matter neutrinos.

B_9 as the gauge boson represents dark matter repulsive force. The condensed baryonic gas at the critical surface density (derived from the acceleration constant a_0 in MOND [13] [14]) induces the creation tensor for dark matter repulsive force to transform dark matter in the region into repulsive dark matter repulsing one another, corresponding to the Farnes' repulsive dark matter. Before the emergence of dark matter repulsive force, dark matter in the CMB was not repulsive.

B_{10} at d = 10 is for the gauge boson for particle-antiparticle asymmetry to provide the slight excess of particle in particle-antiparticle at the Big Bang, while B_8 has particle-antiparticle symmetry. (B_9 emerged long after the Big Bang.) As a result, the excess of particle is α^4 (2.8×10^{-9}) per particle-antiparticle (photon) for the ratio between B_8 and B_{10}. Since baryonic matter is 1/6 of dark matter and baryonic matter from Equation (33), the baryonic matter excess is 4.7×10^{-10} which is in a good agreement with 6×10^{-10} for the ratio of the numbers between baryonic matter and photons in the Big Bang nucleosynthesis [38].

B_{11} is for gravity. F_{11} (8.275×10^{16} GeV) relates to spin 3/2 gravitino, while B_{11} (1.134×10^{19} GeV) relates to spin 2 graviton. In supersymmetry, gravitino and graviton mediate the supersymmetry between fermion and boson in space dimension and gravitation. There are 11 space dimensions in the 11 spacetime dimensional membrane. As a result, the supersymmetry involves 11 F_{11} + B_{11}, which is equal to 1.225×10^{19} GeV in excellent agreement with the Planck mass (1.221×10^{19} GeV) derived from observed gravity as $(\hbar c/G)^{1/2}$ where c is the speed of light, G is the gravitational constant, and ħ is the reduced Planck constant.

Leptons and Quarks

The seven dimensional orbitals are the base for the periodic table of elementary particles [39] [40] to produce the standard model particles and the internal gravity. The periodic table of elementary particles calculates accurately the par-

ticle masses of all leptons, quarks, gauge bosons, the Higgs boson, and the cosmic rays by using only five known constants: the number (seven) of the extra spatial dimensions in the eleven-dimensional membrane, the mass of electron, the masses of Z and W bosons, and the fine structure constant [39]. The calculated masses are in excellent agreements with the observed masses. For examples, the calculated masses of muon, top quark, pion, neutron, and the Higgs boson are 105.55 MeV, 175.4 GeV, 139.54 MeV, 939.43 MeV, and 126 GeV, respectively, in excellent agreements with the observed 105.65 MeV, 172.4 GeV, 139.57 MeV, 939.27 MeV, and 126 GeV, respectively.

4b) the formation of the hidden oscillating dimension number universes from 10D to 5D

The negative-mass 10D particle-antiparticle, the positive-mass external gravity, and the negative-mass external gravity are transformed into the hidden oscillating dimension number universes from 10D to 5D.

$$\begin{aligned} &10\text{D4d} \to 9\text{D5d} \to 9\text{D4d} \to 8\text{D5d} \to 8\text{D4d} \to 7\text{D5d} \\ &\to 7\text{D4d} \to 6\text{D5d} \to 6\text{D4d} \to 5\text{D5d} \to 5\text{D4d} \end{aligned} \tag{39}$$

From Equation (19), under the VSLD transformation and the VSD transformation, the three universes expand through the increasing rest mass and the translation-fractionalization from 10D4d to 5D4d. To the positive-mass 4D universe, the three universes from 10D to 5D are hidden, because as mentioned in Section 3.1.2., particles with different space-time dimensions and different speeds of light are transparent and oblivious to one another to avoid the violation of causality due to differences in the speed of light. During this time, the positive-mass 4D universe expands normally.

5) the transformation of all four universes into the 4D universes

When all four universes become 4D, the three other universes become dark energy as a part of the positive-mass 4D universe.

$$5\text{D4d} \to 4\text{D5d} \to 4\text{D4d} \tag{40}$$

The result is the accelerating expansion. Since the other three universes have no detachment space to produce kinetic energy, dark energy is inert as the inert cosmological constant. As these three universes can take the form of a cosmological constant, the field equations are modified to

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}\left(T^{+}_{\mu\upsilon} + T^{-}_{\mu\upsilon} + C_{\mu\upsilon}\right) \tag{41}$$

where $R_{\mu\nu}$ is the Ricci curvature tensor, R is the scalar curvature, $g_{\mu\nu}$ is the metric tensor, Λ is the cosmological constant, G is Newton's gravitational constant, c is the speed of light in vacuum, $T^{+}_{\mu\upsilon}$ is the positive stress-energy tensor for the positive-mass universe, $T^{-}_{\mu\upsilon}$ is the positive-negative stress-energy tensor for negative-mass universe, negative-mass external gravity, and positive-mass external gravity, and $C_{\mu\nu}$ is the creation tensor for negative-mass universe, negative-mass external gravity, and positive-mass external gravity.

The ratio of the time periods for the transformations from D → D - 1 is pro-

portional to ln of the total number of particles (Equation (22)) to be transformed from D → D - 1 for the exponential growth with time as in Table 3.

The maximum dark energy is 75% for the three out of the four universes, when the spacetime numbers of all particles are 4. The observed % of dark energy is 68.6, and our universe is 13.8 billion-year old [15]. The period of the 5D → 4D is (0.333) (13.8)/((0.333) (68.6/75) + 0.667) = 4.73 billion years, and dark energy as the 5D → 4D started in (4.73) (68.6/75) = 4.33 billon years ago that is in agreement with the observed value of 4.71 ± 0.98 billion years ago when dark energy started to accelerate the cosmic expansion [16].

6) the positive-mass 4D universe and the three hidden oscillating dimension number universes from 5D to 9D

The three oscillating universes from 5D to 10D again become the hidden universes.

$$\begin{aligned} &4\text{D4d} \to 5\text{D4d} \to 5\text{D5d} \to 6\text{D4d} \to 6\text{D5d} \to 7\text{D4d} \to 7\text{D5d} \\ &\to 8\text{D4d} \to 8\text{D5d} \to 9\text{D4d} \to 9\text{D5d} \end{aligned} \tag{42}$$

They contract by the decreasing rest mass and the translation-condensation. The positive-mass 4D universe contracts through gravity. Through symmetry, all four universes contract synchronically and equally.

7) the return to the 10D4d particle-antiparticle universes (step 3)

Eventually, the oscillating universes return to the original 10D. The positive-mass 4D universe reaches the Big Crush to lose all detachment space to become 4D10d, and followed by the deflation to transform into 10D4d. The four universes return to the step 3.

In the positive-energy 4D universe

$$\begin{aligned} &4\text{D} + \text{various d's particles} \xrightarrow{\text{the Big Crush}} 4\text{D10d particles} \xrightarrow{\text{the deflation}} 10\text{D4d particles} \\ &\text{In the other three universes} \\ &9\text{D5d particles} \xrightarrow{\text{VSLD}} 10\text{D4d particles} \end{aligned} \tag{43}$$

From the step 3, the universes can undergo another cycle of the particle-antiparticle universes, or can reverse to the step 2 for the string-antistring dual universes, to the step 1 for the membrane-antimembrane dual universes, and ultimately, to the zero-mass inter-universal void.

4. Protogalaxy and Galaxy Evolutions

As mentioned in the previous section, with positive mass dark matter, the Navarro-

Table 3. The Percentages of the periods in the oscillating dimension number universes. α is the fine structure constant for electromagnetism.

	10D → 9D	9D → 8D	8D → 7D	7D → 6D	6D → 5D	5D → 4D
ratio of total numbers of particles	1	α^{-2}	α^{-4}	α^{-6}	α^{-8}	α^{-10}
ratio of ln (total number of particles)	0	$-2\ln\alpha$	$-4\ln\alpha$	$-6\ln\alpha$	$-8\ln\alpha$	$-10\ln\alpha$
ratio of periods in time	~0	1	2	3	4	5
percentages of periods in time	~0	6.7	13.3	20	26.7	33.3

Frenk-White (NFW) profile [2] derived from standard N-body simulations has the high density dark matter at the core of galaxy, but the high density dark matter at the core of galaxy is not observed [3]. This is known as the core-cusp problem or the cuspy halo problem and is currently unsolved with positive mass dark matter [17].

Another unsolved problem is central supermassive black hole (SMBH) that is located in most massive galaxies. Its origin remains unclear because the greatest difficulty to any theory of SMBH formation has been the observation of very massive ($M \approx 10^9$ $M_\odot$) quasars with SMBH already in place by z ≈ 7, when the universe is just ≈ 800 Myr old long before any stars could grow big or old enough to collapse under their own mass, explode in a supernova and form a black hole [41]. Considerable observational and theoretical evidence shows the viability of massive black hole seeds, formed by the collapse of supermassive stars (SMSs) as a progenitor model for such early, massive accreting black holes [42]. However, it is unclear how such a large SMS remained stable enough without first forming many large stars.

A new unsolved problem is the recent discovery of a cold, massive, rotating disk galaxy (the Wolfe Disk) 1.5 billion years after the Big Bang [43]. Massive disk galaxies like the Milky Way are expected to form at late times in traditional models of galaxy formation [44]. Recent numerical simulations suggest that such galaxies with billion solar masses could form as early as a billion years after the Big Bang through the accretion and merger of cold material [45], but it is not clear how to assemble such a large gas mass while maintaining a relatively stable, rotating disk without breaking into many small galaxies. The Wolfe Disk's large rotational velocity and large content of cold gas remain challenging to reproduce with most numerical simulations.

In this paper, dark matter with and without repulsive force provides an answer to all three unsolved problems, and explains protogalaxy evolution and galaxy evolution. This paper proposes that the condensed baryonic gas at the critical surface density induces the creation tensor for dark matter repulsive force to transform dark matter in the region into repulsive dark matter repulsing one another. Repulsive dark matter removes the DM (dark matter) core-cusp, and provides stabilities for SMS and large protogalaxy before the formation of a large galaxy. Without repulsive dark matter, the universe would have been filled with "train wrecks" without large galaxy structures.

4.1. Protogalaxy Evolution

Protogalaxy evolution consists of six steps: 1) Small Core-Cusp Protogalaxy (BM-DM core cusp-halo), 2) Large Core-Cusp Protogalaxy (BM-DM core cusp-halo), 3) Streaming Protogalaxy (streaming BM core-static BM shell-DM halo), 4) SMBH Protogalaxy (central SMBH-streaming BM core-static BM shell-DM halo). 5) Bulge Protogalaxy (central SMBH-bulge-streaming BM core-static BM shell-DM halo), and 6. Protogalaxy Termination (central SMBH-bulge-combined BM shell-DM halo) as **Figure 5**.

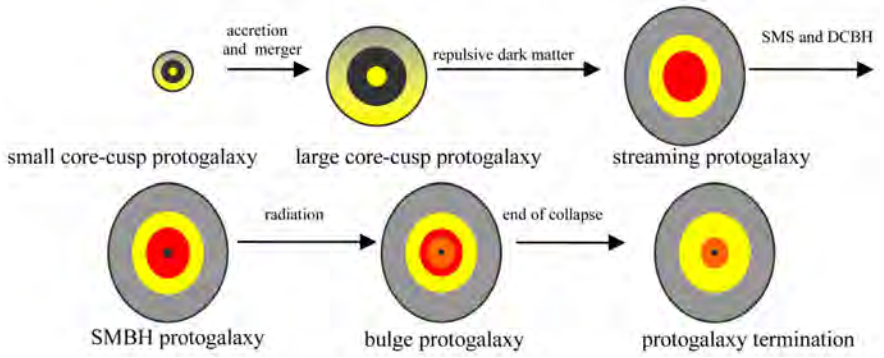


Figure 5. Protogalaxy Evolution: 1) small core-cusp protogalaxy (BM-DM core cusp-halo), 2) large core-cusp protogalaxy (BM-DM core cusp-halo), 3) streaming protogalaxy (streaming BM core-static BM shell-DM halo), 4) SMBH protogalaxy (central SMBH-streaming BM core-static BM shell-DM halo). 5) bulge protogalaxy (central SMBH-bulge-streaming BM core-static BM shell-DM halo), and 6) protogalaxy termination (central SMBH-bulge-combined BM shell-DM halo).

1) The Small Core-Cusp Protogalaxy Step

Quantum fluctuations in the matter distribution were created in the first fraction of a second during an inflationary period. Gravitational instability grew these fluctuations over time. Baryonic matter and dark matter without dark matter repulsive force were initially well mixed. Dark halos emerged as gravitationally bound regions of matter that have decoupled from the cosmic expansion and collapsed. Dark matter haloes were formed as in N-body collisionless dark matter simulations. They followed the NFW profile where the highest density is at the center as the core cusp [2]. Baryonic matter was dragged along by the gravitationally dominant dark matter with the highest density at the center, resulting in small BM-DM core-cusp-halo.

2) The Large Core-Cusp Protogalaxy Step

Dark matter halo structure formation is bottom-up merge with small dark matter haloes forming first. The baryonic gas content can contract together with the dark matter only in dark halos above the cosmological Jeans mass, MJ $\approx 10^4 M_{\odot}[(1 + z)/11]^{3/2}$, in which the gravity of dark matter can overwhelm thermal gas pressure [46]. At z ~ 1100 (400 thousand years old), the temperature drops below ~ 3000K and protons and electrons recombine to form neutral hydrogen. The photons then decouple and travel freely until the present, when they are observed as the CMB (cosmic microwave background) [47] [48]. Following recombination and cooling through the emission of radiation, the neutral gas was freed to fall into the gravitational potential wells of the dark matter, and to condense into the BM core. Unable to dissipate radiation and condense, dark matter stayed outside the BM core to form the DM core cusp-halo, resulting in the large core-cusp protogalaxy consisting of large BM-DM core cusp-halo [49].

3) The Streaming Protogalaxy Step

The condensed baryonic gas at the critical surface density (derived from the acceleration constant a_0 in MOND [13] [14]) induces the creation tensor for

dark matter repulsive force to transform dark matter in the region into repulsive dark matter repulsing one another, corresponding to the Farnes' repulsive dark matter. Dark matter repulsive force in the periodic table of elementary particles for baryonic matter and dark matter (**Table 1**) is a force like the other forces, such as the strong, the weak, electromagnetic, and gravitational forces. The repulsive dark matter particles with dark matter repulsive force in the DM core cusp stream outward to leave a low density dark matter core and to form a halo that surrounds the positive mass galaxy and which extends out to several galactic radii as in the Farnes' unifying theory for repulsive dark matter. The result is the protogalaxy collapse which is like the collapse of a balloon as the air (as dark matter) moving out the balloon. The monolithic protogalaxy collapse forced baryons in the halo region to stream inward to fill the low density core. The result was the streaming protogalaxy with the streaming BM core-static BM shell-DM halo. In some cases, the ejected streaming dark matter jet inevitably carried some baryonic matter out of the protogalaxy. The baryonic matter in the ejected dark matter jet resulted eventually in the dwarf galaxies which are observed as satellite galaxies arranged on a plane as observed by Ibata *et al.* [50].

4) The SMBH Protogalaxy Step

Primordial halos expose to highly supersonic baryon streaming motions [51] [52] which can prevent protogalaxies from forming stars, so the streaming protogalaxies remained stable until they reach masses of 10^7 - 10^8 $M_\odot$ that trigger rapid atomic cooling and catastrophic baryon collapse to start forming SMSs at central infall rates of up to ≈1 $M_\odot$ yr^{-1} in the atomically cooled protogalaxies [53] [54]. The collapse of the SMS at a few 10^5 $M_\odot$ undergoes direct collapse black hole (DCBH) without supernova to produce the central SMBH [55] [56] [57]. The DCBH era lasts from z ≈ 20 to z ≈ 13, or a time period T ≈ 150 Myr. The whole DCBH process stops, when the gas in the atomically cooled halos is photoevaporated by the ambient radiation field [58]. According to the mathematical model of the growth of DCBH in the early universe by Basu and Das [59], the mass growth function of DCBH gains an exponentially large amount of mass in a relatively short amount of time. The rapid growth is capped by the Eddington limit, and for a brief period of time exceeds the Eddington limit before the termination of the DCBH growth. The protogalaxy collapse prevented the fragmentation of baryons during the formation of the SMS as all fragments merged in the core. The result was the SMBH protogalaxy with the SMBH-streaming BM core-static BM shell-DM halo.

5) The Bulge Protogalaxy Step

The incoming streaming baryons during the protogalaxy collapse accelerated the growth of the SMBH. Eventually, the radiation from the SMBH reduced greatly the streaming motion of the incoming streaming baryons, and the SMBH was decoupled from the streaming BM core as described by Joseph Silk, Martin Rees, and Andrew King [60] [61]. Without strong streaming, baryons were fragmented to form normal stars, resulting in the bulge in the streaming BM

core. The result was the bulge protogalaxy with the SMBH-bulge-streaming BM core-static BM shell-DM halo. Since the SMBH and the bulge were formed in the streaming BH core, the mass of the SMBH and the mass of bulge are correlated as found in the studies [62] [63].

6) The Protogalaxy Termination Step

Eventually, without streaming, the streaming BM core and the static BM shell were combined into the combined region, resulting in the SMBH-bulge-combined BM shell (streaming BM core + the static BM shell)-DM halo. The protogalaxy termination means the end of the protogalaxy collapse. The protogalaxy termination step started galaxy evolution when the star formation in the combined region in a large galaxy became active.

4.2. Galaxy Evolution

Elliptical galaxies are mildly flattened, and are mainly supported by the random motions of their stars. Spiral galaxies, on the other hand, have highly flattened disks that are mainly supported by rotation of their stars. Most elliptical and spiral galaxies have both with the ellipsoidal component called the bulge. Irregular galaxies do not have definite shape and the bulge. In this paper, the morphology of galaxies is derived from the monolithic heterogeneous protogalaxy collapse which produced both SMBH and galaxies. Essentially as described in this section, the protogalaxies with the small DM core cusp produced elliptical galaxies, the protogalaxies with the large DM core cusp produced spiral galaxies, and the protogalaxies with the very large DM core cusp produced irregular galaxies.

1) The Formation of Elliptical Galaxy

The progenitor of elliptical galaxy is the protogalaxy with small DM core cusp and large BM shell. The ejection of dark matter from a small DM core cusp resulted in a small empty core which was filled by the streaming baryons to form a small streaming BM core surrounded by a large static BM shell. In terms of star formation, the normal star formation took place only in the large BM shell, while the SMS and DCBH formation took place only in the small streaming BM core. The large amount of baryon gas was depleted by the star formation in the large static BM shell. The stars are in random orbits around the center. The protogalaxy collapse caused the slight flattering of the protogalaxy into elliptical shape whose lengths of major axes are proportional to the relative sizes of the DM core cusp. The early large consumption of gas by the star formation in the large static BM shell depleted large amount of baryon gas, so elliptical galaxies now have very few young stars.

During the early universe, large protogalaxies accreted the surrounding small and medium protogalaxies, and turned them into metal-poor globular clusters without external dark matter haloes [64] in galactic halos. (Some of the accreted protogalaxies had their own accreted protogalaxies.) As a result, the stars in globular clusters in general are older than the stars in the host large galaxies.

2) The Formation of Spiral Galaxy and Barred Spiral Galaxy

The progenitor of spiral galaxy is the protogalaxy with large DM core cusp and small BM shell. The ejection of dark matter from a large DM core cusp resulted in a large empty core which was filled by the streaming baryons to form a large streaming BM core surrounded by a small static BM shell. During the protogalaxy collapse, the streaming baryons initially formed a dense cloud at the center of the streaming BM core. The infalling streaming baryons toward the dense cloud at the center produced the rotational baryon cloud due to the conservation of angular momentum, similar to the rotational cloud during the collapse of the cloud. Eventually, the whole streaming BM core and then the whole protogalaxy became rotational. (In the progenitor of elliptical galaxy, the streaming BM core was too small to rotate the whole protogalaxy.) The stars orbit around the center.

After the termination of protogalaxy collapse, the collapsed protogalaxy consisted of central SMBH-bulge-combined BM shell-DM halo. After the end of the protogalaxy, the protogalaxy underwent differential rotation with the increasing angular speeds toward the center. After few rotations, the protogalaxy turned into the spiral structure consisting of the center disk with the pre-existed bulge-SMBH at the center and the attached spiral arms. The arms consisted of the low density gas regions and the high density gas regions. The high density gas regions hindered the rotational movement of the low baryon gas, so the low baryon gas formed the large high density regions behind the high density gas regions. After few rotations, all large high density regions coalescent into few major arms at the minimum rotational velocities as the high density waves described in the density wave theory by C. C. Lin and Frank Shu [65] [66]. The arms in spiral galaxy are the high density waves. Near the center of a spiral galaxy, the high density wave forms the bar for barred spiral galaxy [67]. The evolution of spiral and barred spiral galaxies is shown in **Figure 6**.

Without the early large depletion of baryon gas in the small BM shell and with the radiation from SMBH and stars to retard the star formation rate, spiral galaxies

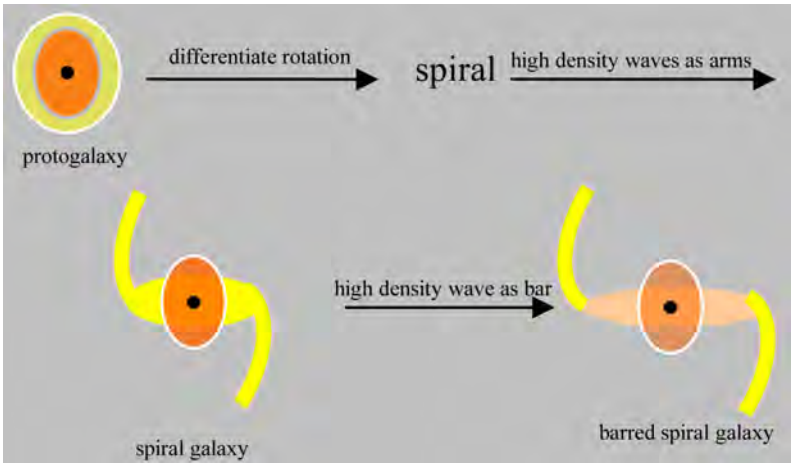


Figure 6. The formations of spiral galaxy and barred spiral galaxy.

and barred spiral galaxies now have many young stars and large amounts of interstellar gas. The stars form in the spiral arms much later than in the bulge and the bar, so they are many young stars in the spiral arms. The lengths of the spiral arms are proportional to the relative sizes of the DM core cusp as the lengths of major axes proportional to the relative sizes of the DM core cusp in elliptical galaxies.

3) The Formation of Irregular Galaxy

If the size of the DM core cusp was very large, the streaming dark matter in the DM core cusp streamed outward, and not enough baryonic matter particles from the BM-DM halo streamed inward to form SMS. Eventually, the protogalaxy became fragmented, resulting in irregular galaxy.

5. Summary and Conclusion

In summary, this paper modifies the Farnes' unifying theory of dark energy and dark matter which are negative-mass to be created continuously from the negative-mass universe in the positive-negative mass universe pair. The first modification explains that observed dark energy is 68.6% greater than 50% for the symmetrical positive-negative mass universe pair. This paper starts with the proposed positive-negative-mass 11D universe pair (without kinetic energy) which is transformed into the positive-negative mass 10D universe pair and the external dual gravities as in the Randall-Sundrum model, resulting in the four equal and separate universes consisting of the positive-mass 10D universe, the positive-mass massive external gravity, the negative-mass 10D universe, and the negative-mass massive external gravity. The positive-mass 10D universe is transformed into 4D universe (home universe) with kinetic energy through the inflation and the Big Bang to create positive-mass dark matter which is five times of positive-mass baryonic matter. The other three universes without kinetic energy oscillate between 10D and 10D through 4D, resulting in the hidden universes when $D > 4$ and dark energy when $D = 4$, which is created continuously to our 4D home universe with the maximum dark energy = 3/4 = 75%. In the second modification to explain dark matter in the CMB, dark matter initially is not repulsive. The condensed baryonic gas at the critical surface density induces dark matter repulsive force to transform dark matter in the region into repulsive dark matter repulsing one another, corresponding to the Farnes' repulsive dark matter. In this way, a galaxy actually started with a dark matter core-cusp without repulsive dark matter. The removal of the core cusp by repulsive dark matter afterward solves many difficult problems in protogalaxy and galaxy evolutions.

Protogalaxy evolution consists of six steps: 1) small core-cusp protogalaxy (BM-DM core cusp-halo), 2) large core-cusp protogalaxy (BM-DM core cusp-halo), 3) streaming protogalaxy (streaming BM core-static BM shell-DM halo), 4) SMBH protogalaxy (central SMBH-streaming BM core-static BM shell-DM halo), 5) bulge protogalaxy (central SMBH-bulge-streaming BM core-static BM

shell-DM halo), and 6) protogalaxy termination (central SMBH-bulge-combined BM shell-DM halo) as Figure 5. In the first two steps, dark matter did not have repulsive force, so photogalaxies were small and large protogalaxies with BM-DM core-cusp-halo. In the step 3, the condensed baryonic gas at the critical surface density induces the creation tensor for dark matter repulsive force to transform dark matter in the region into repulsive dark matter repulsing one another. Through dark matter repulsive force, repulsive dark matter in the core-cusp was ejected from the core, resulting in the protogalaxy collapse which forced baryons in the BM halo to stream inward to fill the nearly empty core to form the streaming BM core. Initially, the density in the streaming BM core was too low to form star. Later, the motions of the streaming baryons prevented forming star until the streaming BM core reached the critical mass of 10^7 - 10^8 $M_\odot$ that triggered rapid atomic cooling and catastrophic baryon collapse, resulting in the supermassive star (SMS) formation in the atomically-cooled streaming BM core. The collapse of the SMS at a few 10^5 $M_\odot$ produced the DCBH for the central SMBH. The bulge was generated in the streaming BM core surrounding the SMBH afterward. During and after the protogalaxy collapse, the star formations produced different types of galaxies (elliptical, spiral, barred spiral, and irregular). Essentially, the protogalaxies with the small DM core produced elliptical galaxies, the protogalaxies with the large DM core produced spiral-barred spiral galaxies, and the protogalaxies with the very large DM core produced irregular galaxies.

According to the theoretical calculation, the calculated percentages of dark energy, dark matter, and baryonic matter are 68.6 (as an input from the observation), 26, and 5.2, respectively, in agreement with observed 68.6, 26.5, and 4.9, respectively, and dark energy started in 4.33 billion years ago in agreement with the observed 4.71 ± 0.98 billion years ago. In conclusion, the modified Farnes' unifying theory reinterprets the Farnes' equations, and is a unifying theory of dark energy, dark matter, and baryonic matter in the positive-negative mass universe pair. The unifying theory explains protogalaxy and galaxy evolutions in agreement with the observations.

Conflicts of Interest

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

References

[1] Farnes, J. (2018) *Astronomy & Astrophysics*, **620**, A92. https://doi.org/10.1051/0004-6361/201832898

[2] Navarro, J., Frenk, C. and White, S. (1997) *The Astrophysical Journal*, **490**, 493-508. https://doi.org/10.1086/304888

[3] Hui, L. (2001) *Physical Review Letters*, **86**, 3467-3470. https://doi.org/10.1103/PhysRevLett.86.3467

[4] Maeder, A., *et al.* (2018) Planck 2018 Results. VI. Cosmological Parameters.

[5] Komatsu, E., *et al.* (2009) *Astrophysics Journal Supplement*, **180**, 330.

[6] Planck Collaboration (2016) *Astronomy Astrophysics*, **594**, A13.

[7] Freese, K. (2017) *International Journal of Modern Physics D*, **26**, Article ID: 1730012. https://doi.org/10.1142/S0218271817300129

[8] Chung, D. and Krasnoholovets, V. (2013) *Journal of Modern Physics*, **4**, 77-84. https://doi.org/10.4236/jmp.2013.47A1009

[9] Chung, D. (2018) *Journal of Modern Physics*, **9**, 2257-2273. https://doi.org/10.4236/jmp.2018.913142

[10] Chung, D. (2019) *Journal of Modern Physics*, **10**, 1310-1341. https://doi.org/10.4236/jmp.2019.1011087

[11] Randall, L. and Sundrum, R. (1999) *Physical Review Letters*, **83**, 3370-3373. https://doi.org/10.1103/PhysRevLett.83.3370

[12] Randall, L. (2005) Warped Passages: Unraveling the Mysteries of the Universe's Hidden Dimensions. Harper Collins, New York.

[13] Milgrom, M. (1989) *Astronomy and Astrophysics*, **211**, 37-40.

[14] Milgrom, M. (2009) *Monthly Notices of the Royal Astronomical Society*, **398**, 1023-1026. https://doi.org/10.1111/j.1365-2966.2009.15255.x

[15] Tanabashi, M., *et al.* (2019) *Physical Review D*, **98**, Article ID: 030001.

[16] Riess, A.G., *et al.* (2004) *Astrophysical Journal*, **607**, 665-687. https://doi.org/10.1086/383612

[17] De Blok, W. (2010) *Advance Astronomy*, **2010**, Article ID: 789293. https://doi.org/10.1155/2010/789293

[18] Holmes, R.B. (2020) *Journal of Modern Physics*, **11**, 648-667. https://doi.org/10.4236/jmp.2020.115042

[19] Chung, D. (2016) *Journal of Modern Physics*, **7**, 642-655. https://doi.org/10.4236/jmp.2016.77064

[20] Chung, D. (2016) *Journal of Modern Physics*, **7**, 1210-1227. https://doi.org/10.4236/jmp.2016.710110

[21] Moffat, J. (1993) *International Journal of Modern Physics D*, **2**, 351-359. https://doi.org/10.1142/S0218271893000246

[22] Albrecht, A. and Magueijo (1999) *Physics Review D*, **59**, Article ID: 043516. https://doi.org/10.1103/PhysRevD.59.043516

[23] Barrow, J. (2003) *Physics Letters B*, **564**, 1-7. https://doi.org/10.1016/S0370-2693(03)00573-2

[24] Weinberg, S. (1989) *Review Modern Physics*, **61**, 1-23. https://doi.org/10.1103/RevModPhys.61.1

[25] Chung, D. (2018) *Journal of Modern Physics*, **9**, 2308-2319. https://doi.org/10.4236/jmp.2018.913146

[26] Gorham, P., *et al.* (2016) *Physical Review Letters*, **117**, Article ID: 071101.

[27] Gorham, P., *et al.* (2018) *Physical Review Letters*, **121**, Article ID: 161102.

[28] Gorham, P., *et al.* (2009) *Astroparticle Physics*, **32**, 10-41.

[29] Bounias, M. and Krasnoholovets, V. (2003) *The International Journal of Systems and Cybernetics*, **32**, 945-975. https://doi.org/10.1108/03684920310483126

[30] Chung, D. (2018) *Journal of Modern Physics*, **9**, 2638-2656. https://doi.org/10.4236/jmp.2018.914164

[31] Aprile, E., *et al.* (2020) Observation of Excess Electronic Recoil Events in XENON1T.

[32] Fujita, J. and Miyazawa, H. (1957) *Progress of Theoretical Physics*, **17**, 360. https://doi.org/10.1143/PTP.17.360

[33] Chung, D. (2016) *Journal of Modern Physics*, **7**, 1150-1159. https://doi.org/10.4236/jmp.2016.710104

[34] Tsui, D., Stormer, H. and Gossard, A. (1982) *Physical Review Letters*, **48**, 1559-1562. https://doi.org/10.1103/PhysRevLett.48.1559

[35] Stormer, H. (1999) *Reviews of Modern Physics*, **71**, 875-889. https://doi.org/10.1103/RevModPhys.71.875

[36] Laughlin, R. (1983) *Physical Review Letters*, **50**, 1395-1398. https://doi.org/10.1103/PhysRevLett.50.1395

[37] Loureiro, A., *et al.* (2019) *Physical Review Letters*, **123**, Article ID: 081301. https://doi.org/10.1103/PhysRevLett.123.081301

[38] Steigman, G. (2007) *Annual Review of Nuclear and Particle Science*, **57**, 463-491. https://doi.org/10.1146/annurev.nucl.56.080805.140437

[39] Chung, D. (2014) *Journal of Modern Physics*, **5**, 1234-1243. https://doi.org/10.4236/jmp.2014.514123

[40] Chung, D. (2016) *Journal of Modern Physics*, **7**, 1591-1606. https://doi.org/10.4236/jmp.2016.712144

[41] Bañados, E., *et al.* (2018) *Nature*, **553**, 473-476. https://doi.org/10.1038/nature25180

[42] Woods, T., *et al.* (2018) Titans of the Early Universe: The Prato Statement on the Origin of the First Supermassive Black Holes.

[43] Neeleman, M., Prochaska, J., Kanekar, N. and Rafelski, M. (2020) *Nature*, **581**, 269-272. https://doi.org/10.1038/s41586-020-2276-y

[44] Fall, S. and Efstathiou, G. (1980) *Monthly Notices of the Royal Astronomical Society*, **193**, 189-206. https://doi.org/10.1093/mnras/193.2.189

[45] Dekel, A., *et al.* (2009) *Nature*, **457**, 451-454. https://doi.org/10.1038/nature07648

[46] Haiman, Z., Thoul, A.A. and Loeb, A. (1996) *Astrophysical Journal*, **464**, 523-528. https://doi.org/10.1086/177343

[47] Spergel, D., *et al.* (2003) *The Astrophysical Journal Supplement Series*, **148**, 175-194. https://doi.org/10.1086/377226

[48] Readhead, A., *et al.* (2004) *Astrophysical Journal*, **609**, 498-512. https://doi.org/10.1086/421105

[49] Yoshida, N., Hosokawa, T. and Omukai, K. (2012) *Progress of Theoretical and Experimental Physics*, **2012**, 01A305. https://doi.org/10.1093/ptep/pts022

[50] Ibata, N., Ibata, R., Famaey, B. and Lewis, G. (2014) *Nature*, **511**, 563-566. https://doi.org/10.1038/nature13481

[51] Hirano, S., Hosokawa, T., Yoshida, N. and Kuiper, R. (2017) *Science*, **357**, 1375-1378. https://doi.org/10.1126/science.aai9119

[52] Schauer, A., Regan, J., Glover, S. and Klessen, R. (2017) *Monthly Notices of the Royal Astronomical Society*, **471**, 4878-4884. https://doi.org/10.1093/mnras/stx1915

[53] Regan, J. and Haehnelt, M. (2009) *Monthly Notices of the Royal Astronomical Society*, **396**, 343-353. https://doi.org/10.1111/j.1365-2966.2009.14579.x

[54] Latif, M., Schleicher, D., Schmidt, W. and Niemeyer, J. (2013) *Monthly Notices of the Royal Astronomical Society*, **432**, 668-678. https://doi.org/10.1093/mnras/stt503

[55] Umeda, H., Hosokawa, T., Omukai, K. and Yoshida, N. (2016) *The Astrophysical*

Journal Letters, **830**, L34. https://doi.org/10.3847/2041-8205/830/2/L34

[56] Woods, T., Heger, A., Whalen, D., Haemmerlé, L. and Klessen, R. (2017) *The Astrophysical Journal Letters*, **842**, L6. https://doi.org/10.3847/2041-8213/aa7412

[57] Haemmerlé, L., Woods, T., Klessen, R., Heger, A. and Whalen, D. (2018) *The Astrophysical Journal Letters*, **853**, L3. https://doi.org/10.3847/2041-8213/aaa462

[58] Yue, B., Ferrara, A., Salvaterra, R., Xu, Y. and Chen, X. (2014) *Monthly Notices of the Royal Astronomical Society*, **440**, 1263-1273. https://doi.org/10.1093/mnras/stu351

[59] Basu, S. and Das, A. (2019) *The Astrophysical Journal Letters*, **879**, L3. https://doi.org/10.3847/2041-8213/ab2646

[60] Silk, J. and Rees, M. (1998) *Astronomy and Astrophysics*, **331**, L1-L4.

[61] King, A. (2003) *The Astrophysical Journal*, **596**, L27-L29. https://doi.org/10.1086/379143

[62] Lasker, R., *et al.* (2017) *The Astrophysical Journal*, **825**, 1. https://doi.org/10.3847/0004-637X/825/1/3

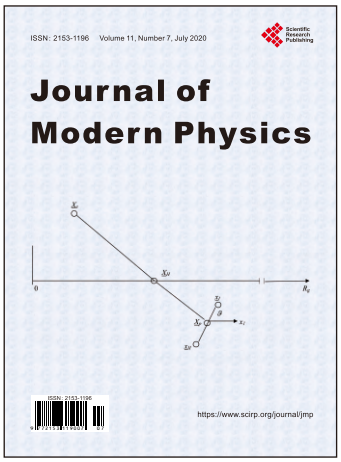
[63] Schramm, M. and Silverman, J. (2013) *The Astrophysical Journal*, **767**, 13. https://doi.org/10.1088/0004-637X/767/1/13

[64] Ibata, R., *et al.* (2011) *The Astrophysical Journal*, **738**, 186. https://doi.org/10.1088/0004-637X/738/2/186

[65] Lin, C. and Shu, F. (1964) *Astrophysical Journal*, **140**, 646-655. https://doi.org/10.1086/147955

[66] Shu, F. (2016) *Annual Review of Astronomy and Astrophysics*, **54**, 667-724. https://doi.org/10.1146/annurev-astro-081915-023426

[67] Bournaud, F. and Combes, F. (2002) *Astronomy and Astrophysics*, **392**, 83-102. https://doi.org/10.1051/0004-6361:20020920



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