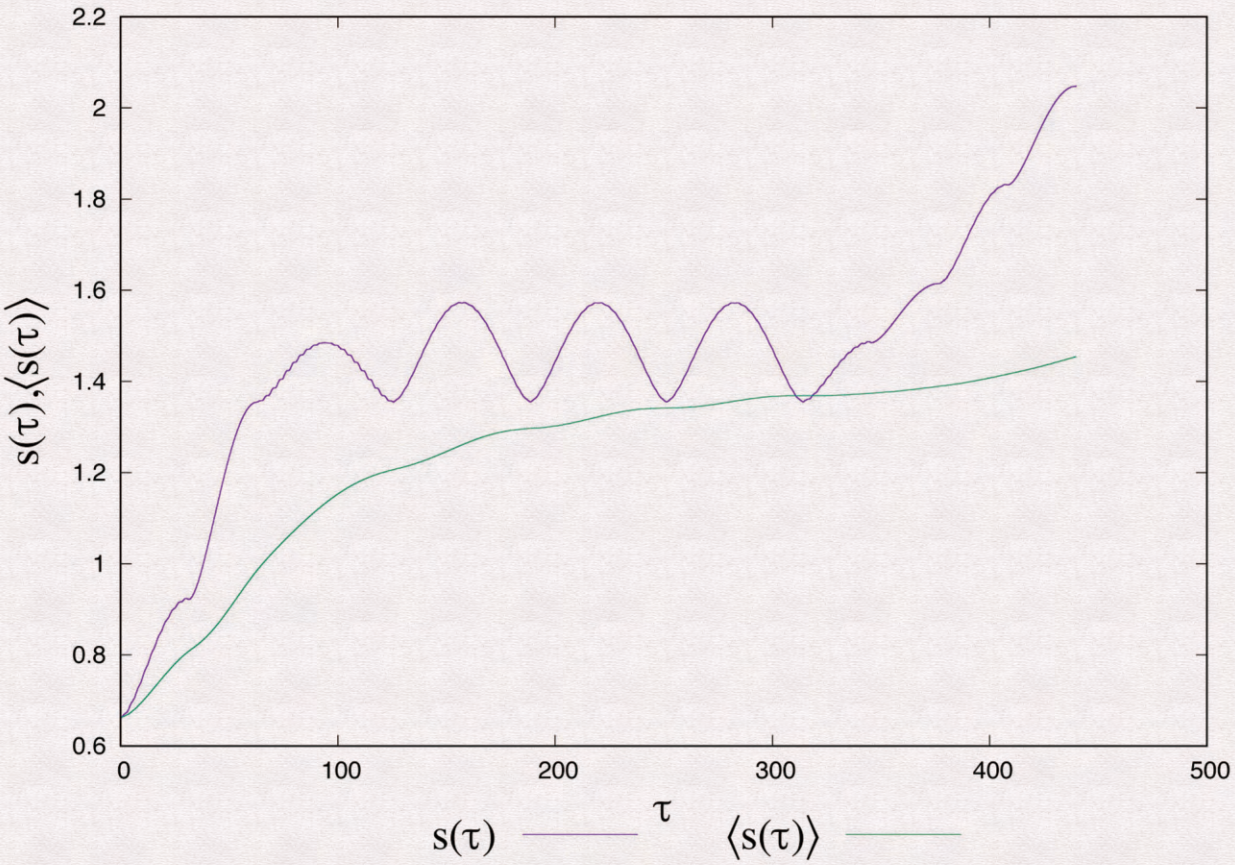
ISSN: 2153-1196 Volume 11, Number 8, August 2020

Scientific Research Publishing

Journal of Modern Physics



ISSN: 2153-1196
9 772153 119007 08

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, 1123-1261
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Table of Contents

Volume 11 Number 8 August 2020

The figure on the front cover is from the article published in Journal of Modern Physics, 2020, Vol. 11, No. 8, pp. 1123-1131 by Gustavo V. López, Jorge Lizarraga.

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, 1123-1131
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

C^{12} - C^{13} Diamond Quantum Computer with Longitudinal rf-Magnetic Field

Gustavo V. López, Jorge Lizarraga

Departamento de Física, CUCEI, Universidad de Guadalajara, Blvd. Marcelino García Barragan y Calzada Olímpica, Guadalajara, Jalisco, Mexico

Email: gulopez@cencar.udg.mx, jorge.a.lizarraga.b@gmail.com

How to cite this paper: López, G.V. and Lizarraga, J. (2020) C^{12} - C^{13} Diamond Quantum Computer with Longitudinal rf-Magnetic Field. *Journal of Modern Physics*, **11**, 1123-1131.
https://doi.org/10.4236/jmp.2020.118070

Received: June 12, 2020
Accepted: July 31, 2020
Published: August 3, 2020

Abstract

In the diamond C^{12} (nuclear spin zero) structure with a linear chain of C^{13} (nuclear spin one half) atoms is applied to a transverse static field with respect of this linear chain, having a gradient along the linear chain, and it is also applied to an rf-magnetic field in a plane with a component in the direction of the static field. It is shown that one qubit rotation, the Controlled-Not (CNOT), the Controlled-Controlled-NoT (CCNOT) quantum gates, and teleportation algorithm can be implemented on this structure using integer multiples of electromagnetic π-pulses. Therefore, a quantum computer can be constructed in this form.

Keywords

Quantum Computer, Diamond, Universal Quantum Gates

1. Introduction

Despite of the great experimental and theoretical efforts, quantum computer realization for 1000 qubits remains being the goal of computer science and physical science. The two main problems faced on this goal are technological limitations (for example, the number of controlled ions trapped, or the maximum number of signal that can be registered by the NMR system), and the decoherence due to interaction of any quantum system with the environment [1] [2] [3] [4]. Several quantum systems have been proposed for quantum computation like ions trapped, atoms trapped, QED cavities, NMR on liquids, chain of nuclear spins one half, superconducting jointions, Diamond VN system, etc. [5]-[11]. One of the most attractive systems (due to their high decoherence time [12]) is nuclear spin systems. One resent proposal of this type [13] is the C^{12} - C^{13} diamond structure for a possible quantum computer, where a chain of C^{13} stable

isotopes of spin one half is embedded periodically along a given line of the diamond lattice. A transverse static field with some gradient along this line, and transversally to it, and rf-magnetic field are applied. The universal CNOT and CCNOT quantum gates are obtained with a single π-pulses (evolution time is $\tau = \pi/\Omega$, being Ω the Rabi's frequency), and Hadamar's gate is obtained with a π/2-pulse (evolution time of the quantum system is $\tau = \pi/2\Omega$). In this paper, we show that even with an rf-magnetic field having a component along the direction of the static magnetic field, it is possible to obtain the CNOT, CCNOT universal quantum gates with a 2π-pulse, the Hadamar's gate with a π-pulse. In addition, the implementation of teleportation algorithm using three qubits is presented with its value of goodness (different of Fidelity [14]) of the reproduction of the algorithm and its associated Boltzmann-Shannon's entropy [15] [16].

2. Analytical Approach

The diamond structure C^{12} (spin zero) with the line of isotopes C^{13} (spin one half) along the x-axis is shown in **Figure 1**. The spin-spin interaction is due to their dipole magnetic moments interaction [17]

$$U = \frac{\mu_0}{4\pi}\frac{3(\boldsymbol{m}_1\cdot\boldsymbol{x})(\boldsymbol{m}_2\cdot\boldsymbol{x})-\boldsymbol{m}_1\cdot\boldsymbol{m}_2}{|\boldsymbol{x}|^3}, \tag{1}$$

where $|\boldsymbol{x}|$ is the distance separation between the magnetic moments $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ which are related with the nuclear spin as

$$\boldsymbol{m}_i = \gamma \boldsymbol{S}_i, \quad i = 1,2. \tag{2}$$

being γ the gyromagnetic ratio ($\gamma \approx 2.675\times10^8$ rad/T·s). Assuming Ising interaction, this energy can be written as

$$U = \frac{2J}{\hbar}S_1^z S_2^z, \quad J = \frac{\mu_0\gamma^2\hbar}{4\pi a^3}, \tag{3}$$

where $\hbar$ is the Planck's constant divided by 2π ($\hbar \approx 1.054571818\times10^{-34}$ J·s) and $a = |\boldsymbol{x}| \sim 10^{-10}$ m is the C^{13} - C^{13} nuclear separation in the diamond crystal. The propose magnetic field is

$$\boldsymbol{B}(x,t) = \left(0, b\cos\theta, -b\cos\theta + B_0(x)\right), \quad \theta = \omega t + \varphi, \tag{4}$$

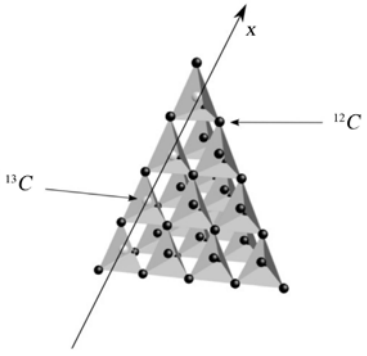


Figure 1. Diamond C^{12} - C^{13}.

where ω, φ and b are the frequency, the phase and the amplitude of the rf-magnetic field, $B_0(x)$ is the static magnetic field transverse to the line of C^{13}'s. The interaction of the C^{13}'s with the magnetic field is given by [17]

$$U = -\sum_{i=1}^{n} \boldsymbol{m}_i \cdot \boldsymbol{B}_i(t), \quad \boldsymbol{B}_i(t) = \boldsymbol{B}(x_i, t), \tag{5}$$

where x_i is the position of the ith-atoms C^{13}. The Hamiltonian of the system considering Ising interaction at first and second neighbor is

$$\hat{H} = U + \frac{2J}{\hbar}\sum_{i=1}^{n-1} S_i^z S_{i+1}^z + \frac{2J'}{\hbar}\sum_{i=1}^{n-2} S_i^z S_{i+2}^z. \tag{6}$$

After rearranging terms, this Hamiltonian can be written as

$$\hat{H} = \hat{H}_0 + \hat{W}, \tag{7}$$

where the operators $\hat{H}_0$ and $\hat{W}$ are defined as

$$\hat{H}_0 = -\sum_{i=1}^{n} \omega_i S_i^z + \frac{2J}{\hbar}\sum_{i=1}^{n-1} S_i^z S_{i+1}^z + \frac{2j'}{\hbar}\sum_{i=1}^{n-2} S_i^z S_{i+2}^z \tag{8}$$

and

$$\hat{W} = \Omega\sum_{k}^{n}\left\{\frac{1}{2i}\left(S_k^+ - S_k^-\right)\cos\theta + S_k^z \sin\theta\right\}, \tag{9}$$

being $\omega_i = \gamma B_0(x_i)$ the Larmore's frequency associated to the ith-atoms C^{13}, $\Omega = \gamma b$ is the so called Rabi's frequency, and S_k^+ and S_k^- are the ascend and descend operators, $S_k^{\pm} = S_k^x \pm S_k^y$. The n-qubits registers $\{|\xi\rangle = |\xi_n, \cdots, \xi_1\rangle\}$ ($\xi_i = 0,1$ for digital notation, and $\xi = 1, \cdots, 2^n$ for decimal notation) form the basis for the Hilbert space of dimensionality 2^n, our qubit is just the spin one half of the nucleus of the atom C^{13}. Now, having the known operations

$$S_i^z |\xi_n, \cdots, \xi_i, \cdots, \xi_1\rangle = \frac{(-1)^{\xi_i}\hbar}{2} |\xi_n, \cdots, \xi_i, \cdots, \xi_1\rangle, \tag{10}$$

$$S_i^+ |\xi_n, \cdots, \xi_i, \cdots, \xi_1\rangle = \hbar\delta_{\xi_i,0} |\xi_n, \cdots, \xi_i + 1, \cdots, \xi_1\rangle, \tag{11}$$

and

$$S_i^- |\xi_n, \cdots, \xi_i, \cdots, \xi_1\rangle = \hbar\delta_{\xi_i,1} |\xi_n, \cdots, \xi_i - 1, \cdots, \xi_1\rangle, \tag{12}$$

being δ_{ij} the Kronecker's delta. The solution of the eigenvalue problems

$$\hat{H}_0 |\xi\rangle = E_\xi |\xi\rangle \tag{13}$$

is given by the eigenvalues

$$E_\xi = -\frac{\hbar}{2}\sum_{i=1}^{n}\omega_i + \frac{\hbar}{2}J\sum_{i=1}^{n-1}(-1)^{\xi_i+\xi_{i+1}} + \frac{\hbar}{2}J'\sum_{i=1}^{n+2}(-1)^{\xi_i+\xi_{i+2}}, \tag{14}$$

and the eigenfunctions are just $\{|\xi\rangle\}$, for $\xi = 1, \cdots, 2^n$. Therefore, proposing a solution of the Schrödinger's equation

$$i\hbar\frac{\partial|\Psi\rangle}{\partial t} = \hat{H}|\Psi\rangle, \tag{15}$$

of the form

$$\left|\Psi(t)\right\rangle = \sum_{\eta=1}^{2^n} \mathrm{e}^{-iE_\eta t/\hbar} D_\eta(t)\left|\eta\right\rangle, \tag{16}$$

we get the equation for the coefficients $D_\xi(t)$ as

$$i\dot{D}_\xi(t) = -i\frac{\Omega\cos\theta}{2}\sum_{\eta=1}^{2^n}\sum_{k=1}^{n} \mathrm{e}^{i\omega_{\xi\eta}t} D_\eta(t)\left\langle\xi\right|S_k^+ - S_k^-\left|\eta\right\rangle + \frac{\Omega(-1)^{\xi_k}}{2} D_\xi(t)\sin\theta, \tag{17}$$

where the frequency $\omega_{\xi\eta}$ has been defined as

$$\omega_{\xi\eta} = \frac{E_\xi - E_\eta}{\hbar}. \tag{18}$$

We can make all system without units by defining the new time evolution as

$$\tau = \omega_0 t, \quad \omega_0 = 2\pi kHz. \tag{19}$$

Then the constant Ω/ω_0, J/ω_0 and ω_i/ω_0 become without units, but the Equation (17) would be exactly of the same form. So, from now on, we will talk about the time evolution in terms of the variable (19), and all the constant must be though as its value times $2\pi kHz$.

3. Quantum Universal Gates

Equations (17) are solved numerically using Runge-Kutta method at fourth order. We made the simulation of the NOT quantum gate (1-qubit), Controlled-Not gate (2-qubits) and Controlled-Controlled-Not gate (3-qubits) to show the realization of a quantum computer with this magnetic field configuration in the C^{12} - C^{13} diamond structure. The parameters used in our simulations are given by

$$\begin{gathered}
\text{1-qubit}: \omega_1 = 100, \quad \Omega = 0.1 \\
\text{2-qubits}: \omega_1 = 100, \quad \omega_2 = 200, \quad \Omega = 0.1, \quad J = 2 \\
\text{3-qubits}: \omega_1 = 100, \quad \omega_2 = 200, \quad \omega_3 = 300, \quad \Omega = 0.1, \quad J = 2, \quad J' = 1/4
\end{gathered} \tag{20}$$

and the initial conditions are

$$\begin{gathered}
\text{1-qubit}: D_1(0) = 1, \quad D_2(0) = 0 \\
\text{2-qubits}: D_1(0) = \frac{1}{\sqrt{16}}, \quad D_2(0) = \sqrt{\frac{2}{16}}, \quad D_3(0) = \sqrt{\frac{5}{16}}, \quad D_4(0) = \sqrt{\frac{8}{16}} \\
\text{3-qubits}: D_1(0) = D_2(0) = \sqrt{\frac{1}{8}}, \quad D_3(0) = D_4(0) = 0, \quad D_5(0) = \sqrt{\frac{3}{8}}, \\
D_7(0) = \sqrt{\frac{2}{8}}, \quad D_8(0) = \frac{1}{\sqrt{8}}
\end{gathered} \tag{21}$$

Figure 2 shows the probabilities $\left|D_i(\tau)\right|^2$ for the system to be in the state $\left|i\right\rangle$ as a function of τ. This curves show the realization of the quantum gates NOT (transition: $\left|0\right\rangle \leftrightarrow \left|1\right\rangle$), CNOT (transition: $\left|10\right\rangle \leftrightarrow \left|11\right\rangle$), and CCNOT (transition: $\left|110\right\rangle \leftrightarrow \left|111\right\rangle$) using a 2π-pulse ($\tau = 2\pi/\Omega$). The probabilities of the other states (CNOT and CCNOT) remain constant within and error of the order 10^{-5}. In addition, we must observe that a superposition of states

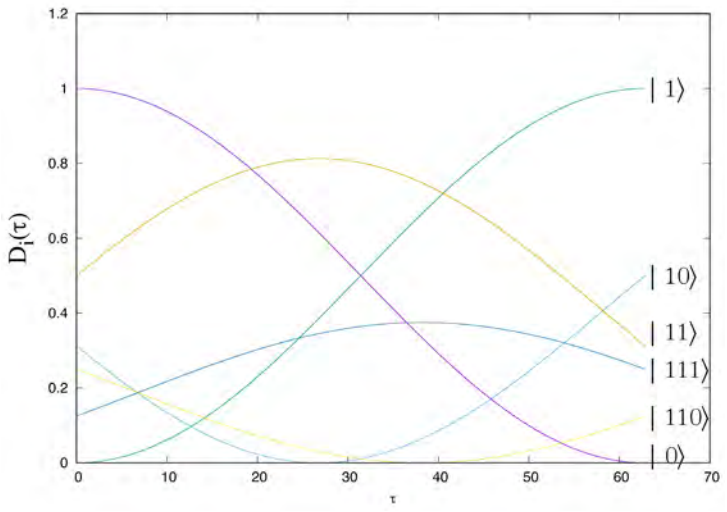


Figure 2. Transitions: $|0\rangle \leftrightarrow |1\rangle, |01\rangle \leftrightarrow |11\rangle, |110\rangle \leftrightarrow |111\rangle$.

(Hadamar quantum gate) is gotten for a time evolution of $\tau = \pi/\Omega$ (π-pulse), that is the states of the form

$$\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \quad \text{(1-qubit)} \tag{22}$$

where some phases could be as factors on each state. We want to mention that if the initial conditions for CNOT and CCNOT are $D_3(0) = 1$ and $D_7(0) = 1$ and all other coefficients are zero, the superposition of states $|10\rangle + |11\rangle$ and $|110\rangle + |111\rangle$ are obtained too.

4. Realization of Teleportation Algorithm

Quantum teleportation [18] is one of the most fundamental elements in quantum computation and quantum information which deserves its realization for any quantum computer proposal. Therefore, this algorithm will be studied in this section with three qubits configuration, where the first state will represent Bob, the second state will be Alice and the third state will be the teleported state. From the Alice location, we want to transport this state (in general, an unknown state even by Alice) to Bob place. If the state we want to transport is

$$|\psi\rangle = c_1|0\rangle + c_2|1\rangle, \quad c_1 = \sqrt{\frac{3}{8}}, \quad c_2 = \sqrt{\frac{5}{8}}, \tag{23}$$

the initial state of our 3-qubits system is then

$$|\Psi_0\rangle = |\psi\rangle \otimes |0\rangle \otimes |0\rangle = c_1|000\rangle + c_2|100\rangle. \tag{24}$$

The "ideal" (mathematical) teleportation algorithm is represented by quantum circuit shown on **Figure 3** (going from left to right), or written in term of operators we have

$$T = \left(H_3 \cdot CN_{32}\right) \cdot \left(CN_{21} \cdot H_2\right), \tag{25}$$

where CN_{ij} is the CNOT operation between the qubit "i" and the cubit "j"

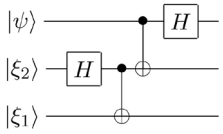


Figure 3. Teleportation quantum circuit.

$$CN_{ij}\left|\xi_n,\cdots,\xi_i,\cdots,\xi_j,\cdots,\xi_1\right\rangle=\left|\xi_n,\cdots,\xi_i,\cdots,\xi_j\oplus\xi_i,\cdots,\xi_1\right\rangle, \tag{26}$$

and H is the Hadamar operator (1-qubit). The first term, $\left(CN_{21}\cdot H_2\right)$, makes the entanglement between Alice(second qubit) and Bob (first qubit), and the term $\left(H_3\cdot CN_{32}\right)$ represents the Alice operations to produce the teleportation. The result of the application of the operator (25) to the initial state (24) can be written as

$$T\left|\Psi_0\right\rangle=\frac{1}{2}\left\{\left|00\right\rangle\otimes\left|\psi\right\rangle+\left|01\right\rangle\otimes N_3\left|\psi\right\rangle+\left|10\right\rangle\otimes Z_3\left|\psi\right\rangle+\left|11\right\rangle\otimes N_3Z_3\left|\psi\right\rangle\right\}, \tag{27}$$

where N_3 (NOT gate), Z_3 and N_3Z_3 act in the following way

$$N_3\left|\psi\right\rangle=c_1\left|1\right\rangle+c_2\left|0\right\rangle,\quad Z_3\left|\psi\right\rangle=c_1\left|0\right\rangle-c_1\left|1\right\rangle,\quad N_3Z_3\left|\psi\right\rangle=c_2\left|0\right\rangle-c_1\left|1\right\rangle. \tag{28}$$

Thus, the state $\left|\psi\right\rangle$ has been teleported to Bob's place, and this has occurred instantly (independently how separated are Alice and Bob). Of course, Alice needs to notify (at the speed of light or less) to Bob about the operation he needs to do in order to get the state $\left|\psi\right\rangle$ ($N_3^2=Z_3^2=I$, the identity operator). Notice then that at the end of the algorithm, we have four states with a probability $\left|c_1\right|^2/4$ and four states with probability $\left|c_2\right|^2/4$. However, in our non-ideal system, we need to apply the correct type of pulse with the desired frequency transition in order to get the state what we want. Let us denote by

$$R_{ij}\left(\tau\right), \tag{29}$$

the pulse of length τ done in the system with a resonant frequency $\omega=\left(E_i-E_j\right)/\hbar$. Our non-ideal teleportation operator is given by

$$\tilde{T}=\left[R_{26}\left(\pi\right)R_{37}\left(\pi\right)R_{78}\left(\pi\right)R_{15}\left(\pi\right)R_{78}\left(2\pi\right)R_{57}\left(2\pi\right)\right]\left[R_{78}\left(2\pi\right)R_{34}\left(2\pi\right)R_{57}\left(\pi\right)R_{13}\left(\pi\right)\right], \tag{30}$$

which represents π and 2π pulses where the Rabi's frequency has been omitted. The last square parenthesis represents the entanglement, and the other one represents the teleportation. The result of this application to the initial state (24) can be seen **Figure 4**, where we have gotten four states with the probability 3/32 and four states with probability 5/32.

The goodness of the reproducibility of the teleportation algorithm can be measured by comparing the probabilities of the states between the real case with the ideal case at the end of the algorithm, and this can be done through the quantity

$$G=\sum_{i=1}^{8}\left|\left|D_i^{(ideal)}\right|^2-\left|D_i^{(real)}\right|^2\right|, \tag{31}$$

which goes from zero (excellent reproducibility) to one (non reproducibility at all). In our case, the value obtained is $G=7.4\times10^{-4}$, meaning that we have a

very good reproduction of the teleportation algorithm.

On the other hand, it is of interest to see how the information of the system is lost during the teleportation algorithm since starting with 2 states we finish with 8 states. Therefore, there must be an increasing on Boltzmann-Shannon entropy and its average value,

$$S(\tau)=-\sum_{i=1}^{8}\left|D_i(\tau)\right|^2\ln\left|D_i(\tau)\right|^2,\quad \left\langle S(T)\right\rangle=\frac{1}{T}\int_0^T S(\tau)\mathrm{d}\tau. \tag{32}$$

Of course, the maximum value this entropy can have is went all the probabilities are the same (1/8), $S_{\max}=\ln 8\approx 2.08$. Figure 5 shows the evolution of the entropy and its average value, and as we expected, there is an increasing on the entropy and its average value due to the increasing of the number of states involved in the quantum dynamics of teleportation algorithm. The oscillations shown on this figure are due to CNOT and CCNOT operations which change the states but do increas the number of states. The rapidly increasing on the entropy is due to Hadamar's operators which increases the number of states involved in the quantum dynamics.

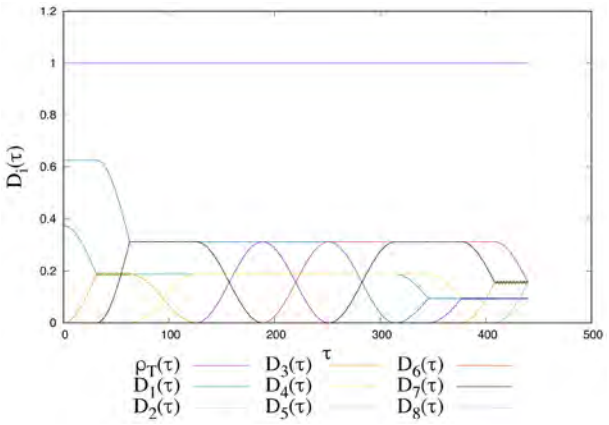


Figure 4. Teleportation algorithm and total probability $P_i(\tau)$.

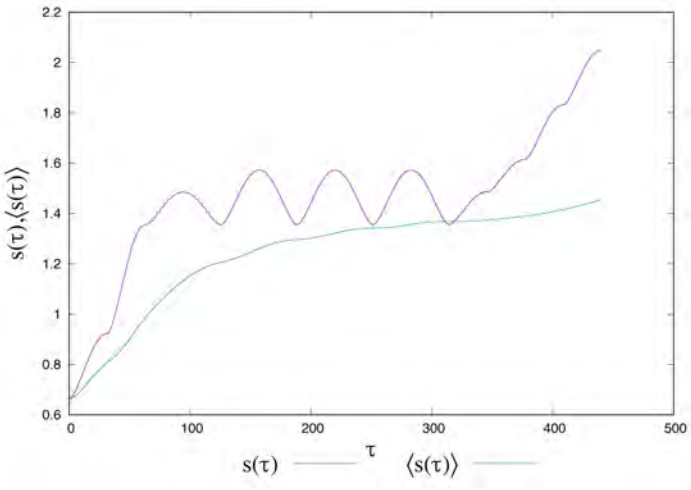


Figure 5. Boltzman-Shannon entropy for the teleportation algorithm.

5. Conclusions

We have shown that a solid state quantum computer can be realized using C^{12} - C^{13} structure in a diamond with even a longitudinal rf-magnetic field (one of its components is along the static field). The universal CNOT and CCNOT quantum gates and the one qubit rotation were well established on the simulations with multiples of a π-pulse. The teleportation algorithm was also implemented using these quantum gates in a 3-qubits configuration and using 10 pulses, and its associated Boltzmann-Shannon's entropy was determined.

Conflicts of Interest

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

References

[1] Warren, W.S. (1997) *Science*, **277**, 1688. https://doi.org/10.1126/science.277.5332.1688

[2] Vandersypen, L.M.L., Steffe, M., Breyta, G., Yannoni, C.S., Sherwood, M.H. and Chuang, I.L. (2001) *Nature*, **414**, 883. https://doi.org/10.1038/414883a

[3] Holzschelter, M.H. (2002) *Los Alamos Science*, **27**, 264.

[4] Monroe, C. and Kim, J. (2013) *Science*, **339**, 1164. https://doi.org/10.1126/science.1231298

[5] Walter, H., Varcoe, B.T.H., Englert, B.G. and Becker, T. (2006) *Reports on Progress in Physics*, **69**, 1325. https://doi.org/10.1088/0034-4885/69/5/R02

[6] Jaksch, D., Cirac, J.I., Zoller, P., Rolston, S.L., Coté, R. and Lukin, M.D. (2000) *Physical Review Letters*, **85**, 2208. https://doi.org/10.1103/PhysRevLett.85.2208

[7] Younge, K.C., Knuffman, B., Anderson, S.E. and Raithel, G. (2010) *Physical Review Letters*, **104**, Article ID: 173001. https://doi.org/10.1103/PhysRevLett.104.173001

[8] Chiorescu, I., Nakamura, Y., Harmans, C.J.P.M. and Mooij, J.E. (2003) *Science*, **299**, 1869. https://doi.org/10.1126/science.1081045

[9] Kitaev, A.Yu. (2003) *Annals of Physics*, **303**, 2-30. https://doi.org/10.1016/S0003-4916(02)00018-0

[10] Childress, L. and Hanson, R. (2013) *MRS Bulletin*, **38**, 134-138. https://doi.org/10.1557/mrs.2013.20

[11] Berman, G.P., Kamenev, D.I., Doolen, D.D., López, G.V. and Tsifrinovich, V.I. (2002) *Contemporary Mathematics*, **305**, 13. https://doi.org/10.1090/conm/305/05213

[12] Watanabe, S. and Sasaki, S. (2003) *Japanese Journal of Applied Physics*, **42**, L1350. https://doi.org/10.1143/JJAP.42.L1350

[13] López, G.V. (2014) *Journal of Modern Physics*, **5**, 55. https://doi.org/10.4236/jmp.2014.51009

[14] Fuchs, C.A. and Caves, C.M. (1994) *Physical Review Letters*, **73**, 3047. https://doi.org/10.1103/PhysRevLett.73.3047

[15] Boltzmann, L. (1866) *Wiener Berichte*, **53**, 195.

[16] Shannon, C.E. and Weaver, W. (1949) The Mathematical Theory of Communica-

tion. University Illinois Press, Champaign.

[17] Jackson, J.D. (1999) Classical Electrodynamics. Third Edition, John Wiley and Sons, Inc., Hoboken, Chapter 5.6.

[18] Nielsen, M.A. and Chuang, I.L. (2000) Quantum Computation and Quantum Information. Cambridge University Press, Cambridge.

Journal of Modern Physics, 2020, 11, 1132-1156
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Generic Description Model for the Structure of Atomic Nucleus with New Interpretation of the Strong Forces

Hongguang Yang, Weidong Yang

Munich, Germany
Email: hyang2013@gmail.com, wdyconf@gmail.com

How to cite this paper: Yang, H.G. and Yang, W.D. (2020) A Generic Description Model for the Structure of Atomic Nucleus with New Interpretation of the Strong Forces. *Journal of Modern Physics*, **11**, 1132-1156.
https://doi.org/10.4236/jmp.2020.118071

Received: June 24, 2020
Accepted: August 3, 2020
Published: August 6, 2020

Abstract

The present investigation is motivated by finding and developing an easily understandable solution in the context of unified quantum and gravitational theories. Model-based methods are applied, with emphasis on structural descriptions by introducing some strong hypotheses. A subset of the introduced hypotheses led to a surprising understanding of the internal structure and construction of quarks, neutrons, protons and more complex atomic nuclei. The research work therefore focused mainly on the model-based interpretation of subatomic processes. The results obtained so far and presented in this paper are new. They consist of a generic description model for the structure of atomic nuclei. This model contains two important structural links that originate from the initial phase of the cosmological big bang. They hold atomic parts together and are involved in many known nuclear fusion and fission processes. Modifications of them, including the electron-positron annihilation process, are necessary and will be described. A new interpretation of the strong forces from the Standard Model is possible and will be given. In addition, the formation processes for electron and positron particles are considered. Based on the structural relationships, a deeper understanding of matter transformations (transmutations), early cosmological processes and dark matter has been achieved. All challenges of this work are the logical conclusions from the used hypotheses on two structural links. They need to be further investigated and verified by theoretical and experimental works. The postulated particle in this paper, as accompanying product in the electron-positron annihilation, will play a major role for the future investigations.

Keywords

Atomic Kernel Structure, Reinterpretation of Strong Forces, Kernel Structure of Hydrogen and Its Isotopes, Kernel Structure of Helium and Its Isotope,

Stellar Nuclear Fusion, Nuclear Fission, Double Space Link, Triple Space Link, Matter Lifecycle in the Context of Cosmology

1. Introduction

The commonly used methods for the analysis of subatomic processes from the well-known Standard Model and the Quantum Chromodynamics Model QCD focus mainly on elementary particles and their interactions with each other (Ref. [1] [2] [3] [4]). Recent experimental data provide different aspects of our understanding of strong forces for the atom core structure (Ref. [5]-[10]). More systematic research on the structural make-up of atomic nuclei, e.g. how neutrons and protons are geometrically arranged within a nucleus, can help to improve the analysis and understanding of subatomic processes and macro-physical aspects.

The generic description model, the novelty of this work—we call it the YY model—aims at this question and can help to achieve improvements. The starting point, unlike what we usually do to explore the quantum nature of atomic and subatomic physics, is the introduction of two essential structural instances that combine all materialized quarks to form an atomic nucleus. These are the "Pairing Space Link PSL" and the "Triple Space Link TSL". The first, PSL, allows a new description of quarks, neutrons and protons in a deterministic spatial structure (Sections 2 and 3). The second, TSL, extends the design scheme to build complex atomic nuclei and isotopes, as applied for deuterium, tritium and helium (Sections 4, 5, 6, 7 and 8).

"Yin" and "Yang", which represent conceptual units of one third of an electric charge, negative and positive respectively, are defined. Hypothetically, a pair of them forms a PSL, while three PSLs form a TSL. The common node must always be "triple charged" for a complete positive or negative electrical unit. In addition, and as a rule, the numbers of Yin's and Yang's must be preserved in all nuclear transformations.

The hypotheses are not proven. However, their effects are clear and easy to understand. The YY model based on PSL/TSL allows a deterministic view of all the building quarks within or around a nucleus and their disassembly and reassembly in all nuclear transformation scenarios. Electric charges also include the inner structural instances. The internal charge balance effect (ICB, Sections 8 and 11), which effectively replaces, or stands for the QCD mechanism, allows a new interpretation of the strong forces of the Standard Model. As a result, the interaction mechanism of QCD could be simplified or reduced to those for PSL and TSL. The logical conclusion is a well-defined modification of our understanding of the electron, positron and their annihilation process (Sections 9 and 10).

Taking into account known facts about kernel construction, design rules for the YY model are introduced (Section 11). We were able to explain some

well-researched processes of stellar nuclear fusion and predict a few unknown aspects that play a role in these processes (Section 12). Furthermore, the involvement of TSLs in nuclear fission is considered (Sections 13).

Using the YY model, more structural details will be visible for an atomic nucleus than is the case with the conventional description. It allows very fine considerations of the spatial distribution of protons and neutrons within or around an atomic nucleus—as we will see, this distribution is a matter of their up and down quarks.

In connection with TSL, a new "Y-particle" is postulated, which plays an important role in the energy-matter life cycle around nuclear fusion and fission processes, in electron-positron annihilation and in the formation of matter during the inflation phase of the big bang. Based on the predicted origin of TSL, a rough description of some macro-physical processes in the cosmological evolution is given (Sections 14). A possible role of TSL for dark matter is presented (Section 15).

Following a model-driven approach—a methodology widely used in computer science—mathematical formulations are completely omitted from the investigations in this paper. In addition, energy balances are not formulated in the transformations. Many different aspects that can be derived from this new mechanism have to be verified afterwards. We hope to gain the confidence of many other physicists in the YY model. They could extend the core idea of this model in conjunction with quantum field theories and carry out experimental verifications, as well as linking it with existing big bang theories for space and cosmology, especially with the study of dark matter.

2. Extended Description for up and down Quark

The starting point is the conventional description of the elementary particles, up quark and down quark, in considering their essential properties: An up quark possesses 2/3 positive of an electrical charge, whereas a down quark has 1/3 negative of an electrical charge. Consequently, two basic building units are elaborated as following:

- "Yin" unit corresponding a third of negative charge, symbolized by "–"
- "Yang" unit corresponding a third of positive charge, symbolized by "+"

These two units, supposed its physical existence, attract each other, and they keep from each other too, so that there is a stable state by "pairing" of them, building a space link, thus called by "pairing space link PSL", as described in Figure 1.

This construction can be expressed by {–|+}. As will be seen later, a space link by pairing is a part of the basic concept of YY model to materialize the energy-matter-space interactions assigned to the pairing. The curly bracket "{ }" and the pipe sign "|" symbolize this fact.

Figure 1. Yin-Yang paring for building a physical space link.

By using Yin, Yang and PSL, quarks will be described as following:

- A down quark contains and materializes a yin-unit, expressed as symbol (−)
- An up quark contains and materializes two yang-units and a space link, expressed by (++{+|−})

Drawing the geometric forms of up quark and down quark results in Figure 2.

Parenthesis "( )" in the expressions and shadows in the Figures symbolize the other form of the energy-matter-space interactions by the quarks.

3. Extended Description for Neutron and Proton

Since a neutron consists of two down and one up quarks and a proton consists of one down and two up quarks (Ref. [2]), the YY model results in the description of neutron and proton in Figure 3.

This re-modeling of neutron and proton reveals more details of their inner structure. The net summation of the positive and negative electrical charges, respectively for a proton and for a neutron corresponds to the classic description. The only difference comes from the space links: they hold the nucleus together.

The following expressions for neutron and proton are equivalent descriptions in Figure 3:

- Neutron: (++{+|−})(−)(−)
- Proton: (++{+|−})(−)({−|+}++)

The meta model for the atomic nucleus presented here could lead to concrete predictions for the interaction behavior of neutrons and protons, with each other or with electric fields. For example (here only as a purely theoretical consideration):

- ✓ A neutron has a space orientation respective to the inner charge distribution, namely the axis from the positive to negative pool, as showed in Figure 4(a), though the whole system is electrically neutral. From this point of view, a neutron, released in an electrical field, should orient its positive end towards the negative side of the electrical field, Figure 5(a).

Figure 2. Geometric forms of up quark and down quark.

Figure 3. Description model for neutron and proton.

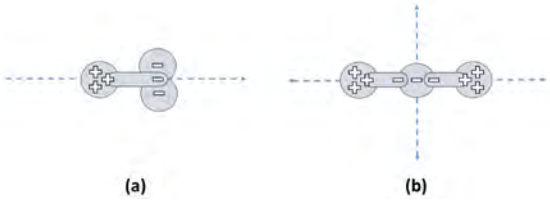


Figure 4. Rotation axes of neutron and proton.

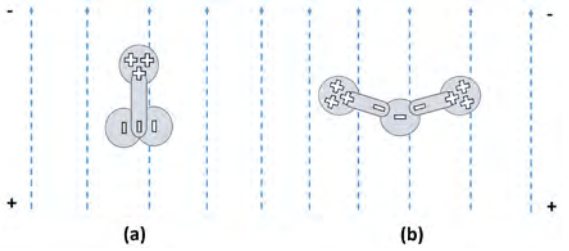


Figure 5. Orientation of neutron and proton in an electrical field.

✓ A proton has a rotation symmetry respective to the axis through the charged poles "positive - negative - positive"; see Figure 4(b). The axis orientation from left to right does not differ from right to left. From this point of view, a proton, released in an electrical field, should tend to orient its symmetry axis perpendicularly to the electrical field, Figure 5(b). A possible electrostatic deformation of the proton structure is also indicated there.

Of course, more research is needed to prove the significance of the effects for an observability, for example whether they can be measured macroscopically as polarization.

By given the proton model, as described above, further descriptions for atomic kernels of the hydrogen isotopes deuterium H-2 and tritium H-3 (Ref. [11]) will continue.

4. Triple Space Link and Its Role for the Atomic Kernel

To model complex kernels, a new construct "triple space link (TSL)" is necessary:

- Three PSLs can bind themselves together on their yang units. Their common node corresponds a positive electrical charge unit ("triple charging"), see Figure 6(a);
- Similarly, three PSLs can bind themselves together on their yin units to build a negative electrical charge unit; see Figure 6(b).

It is still unclear, how three basic units of a common sign are linked in a node for charging from a physical perspective of view. Furthermore, triple charging is supposed to make an essential difference between the positive and negative nodes in the intensity of binding, as indicated in Figures (a) and (b). The binding on the positive node is strong, whereas the connectivity on the negative node is weak (displayed as gaps) and will be re-connected during the nuclear transformations:

- There is an asymmetry between positive charged TSL and negative charged TSL: Whereas a positive charged TSL is pre-created and bound strongly, a negative charged TSL is a mixture of yin pools from positive charged TSLs, of up quarks or of a down quarks. Figure 7 gives all possible combinations for negative charged TSLs, inclusive neutron and proton.

This asymmetry corresponds to the asymmetry between down and up quarks, which will be considered again in a very later section concerning the origins of

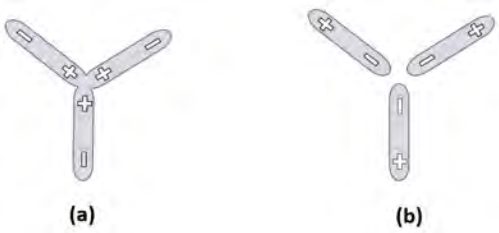


Figure 6. Triple space links, positive and negative charged.

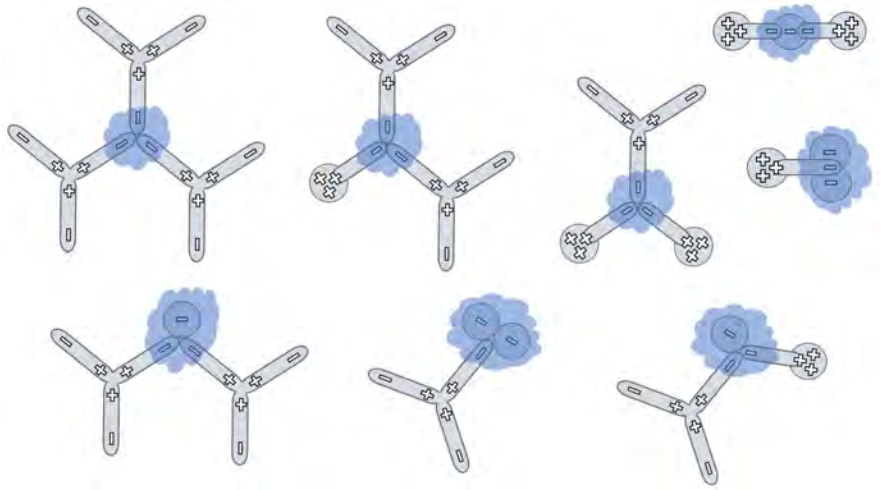


Figure 7. All possible combinations for negative charged TSLs.

cosmology. Due to the asymmetry, speaking of TSL in all following sections we mean exclusively the instances with a positive central node, as drawn in Figure 6(a). Following rules will apply:

- Pairing space link (PSL) can be embedded in an up quark, as described in Section 2;
- Triple pairing space links can be merged on the common positive ends (yang-ends) to build a triple space link (TSL), Figure 6 and Figure 7;
- Atomic kernel building is a combination of up quarks by their yin-ends, TSLs by their yin-ends, and down quarks by themselves. Yang-ends stay untouched.
- Conservation of Yin's and Yang's: In all transformation processes, the whole number of Yin's and the whole number of Yang's remain unchanged.

5. Hydrogen Atomic Kernel H-2 and Re-Interpretation for Strong Force

The starting point is formally going to disassemble an existing proton and an existing neutron and get all their building quarks, see Figure 8. Parts are colored for easy identifying before and after the transformation.

The next step is to reassemble them to a single atomic kernel for building a deuterium (Ref. [11]), with a positive electrical charge unit and a mass unit of two (namely the mass summation of a proton and of a neutron). There are many

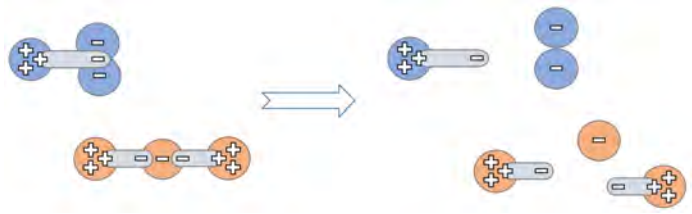


Figure 8. Disassembling a proton and a neutron to their building quarks.

possible ways for reassembling the given quarks, turning out different (quark-) configurations. Only certain configurations of down and up quarks are able to satisfy the charge and mass restrictions for a deuterium.

Starting from rules for quark configurations (at the end of the last section), more rules will be introduced in later sections, while considered atomic kernels become more complex. In the case for reassembling a deuterium, a new triple space link TSL is needed, which combines all given down and up quarks for getting a consistent deuterium atomic kernel, as described in Figure 9.

The upcoming of a TSL in the process of reassembling is a very important point and will be treated in further discussions later.

In consideration of the electrical charges of the whole aggregate above: The net summation of the positive electrical charges is three time 2/3 (=> two positive charge units) and the net summation of the negative electrical charges is three time 1/3 (=> one negative). The net balance of all electrical charges turns out one positive.

In consideration of the mass of the whole aggregate above: The net summation of all three up and all three down quarks results a complete mass of two atomic units, namely three times a third from up quarks, plus three time a third from down-quarks ($3 \times \{1/3\} + 3 \times \{1/3\} = 6/3 = 2$).

The one positive triple link and one negative triple link in the interior of the aggregate are equilibrated in charges.

The strong forces of standard model holding the proton and the neutron together are re-interpreted: The interior negative and positive triple charges of space links hold the surrounding charged up and down quarks together—the holding forces are short-ranged electrostatic forces themselves. QCD based on dynamic color charges of quarks, anti-quarks and gluons (Ref. [12]) is not necessarily used here.

Please also note the drawing of the atomic aggregate in a two-dimensional plane: For a three-dimensional description in the real world, electrostatic attraction and repulsion of up and down quarks around the entire aggregate must be taken into account. In the case of Figure 9, the double down quarks at the top (or/and the double up quarks on the left) must be arranged at least perpendicular to the plane to compensate for electrostatic forces.

6. Hydrogen Atomic Kernel H-3

By applying the reassembling schema used above, it is easy to get a consistent

atomic kernel model for the tritium (Ref. [11]). Firstly, an existing deuterium and an additional neutron are disassembled in Figure 10.

Thereafter, all parts are reassembled together, by adding a new triple space link TSL, resulted as a tritium kernel in Figure 11(a). It is also possible to get a reassembled alternatively model in Figure 11(b).

Please also note the upcoming of a TSL in the process of reassembling.

In consideration of the electrical charges in the aggregate above: The net summation of the positive electrical charges is $4 \times \{2/3\}$ (four up quarks times 2/3) and the net summation of the negative electrical charges is $5 \times \{1/3\}$ (five down quarks times 1/3). The net balance of all electrical charges turns out one positive ($4 \times \{2/3\} - 5 \times \{1/3\} = 3/3 = 1$).

In consideration of the mass in the aggregate above: The net summation of all four up and all five down quarks results a complete mass of three atomic units (four times a third from up quarks, plus five time a third from down-quarks $4 \times \{1/3\} + 5 \times \{1/3\} = 9/3 = 3$).

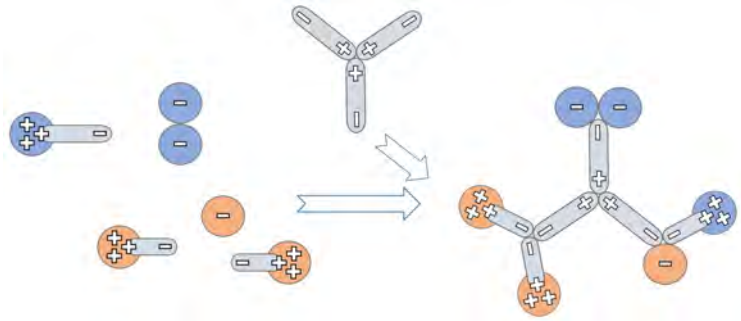


Figure 9. Reassembling the given quarks to a deuterium by adding a TSL.

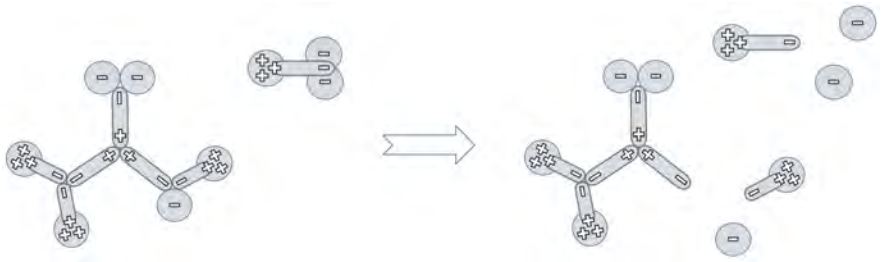


Figure 10. Disassembling a deuterium and a neutron.

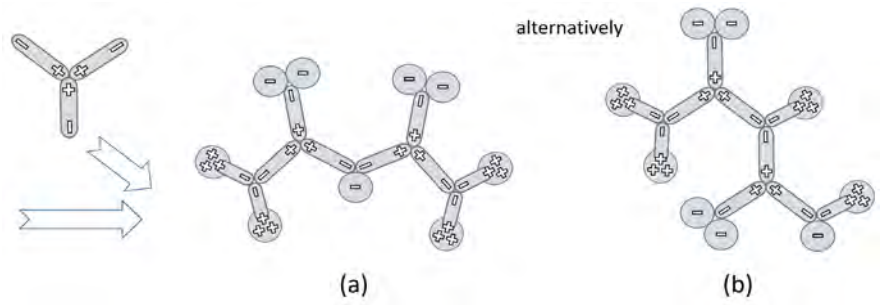


Figure 11. Two alternative tritium kernel models reassembled by adding a TSL.

The two positive triple nodes and two negative triple nodes in the interior of the aggregate are equilibrated in charges.

Similarly, the classic strong forces holding the proton and the neutrons are re-interpreted by links of different sorts. This is the same mechanism for holding one proton and one neutron together.

7. Helium Kernel He-4 and Its Isotope He-3

This section will describe the helium atomic kernel He-4 in YY model, constructed by two protons and two neutrons. Furthermore, one of its isotopes He-3 will also be considered, with two protons plus one neutron.

In the case of He-4, a formal reassembling from two H-2 atomic kernels is done. Figure 12 firstly disassembles two H-2 atom kernels respectively.

By adding one triple space link TSL and reassembling them all, a stable helium kernel He-4 is resulted, see Figure 13.

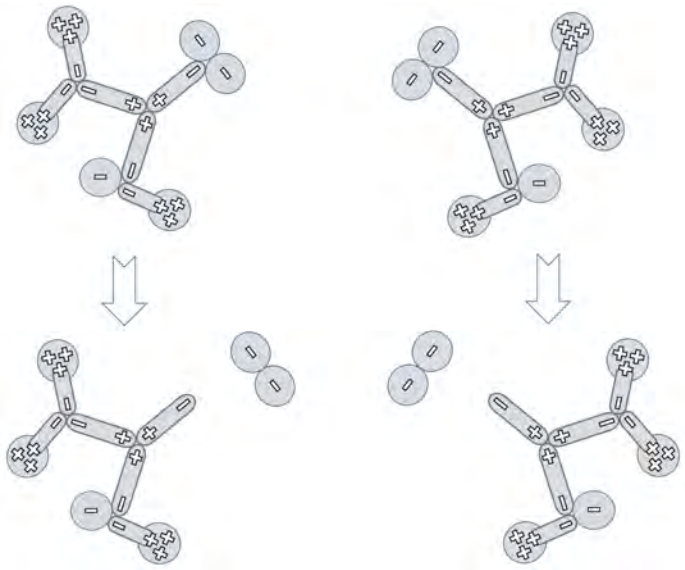


Figure 12. Disassembling two hydrogen atomic kernels H-2.

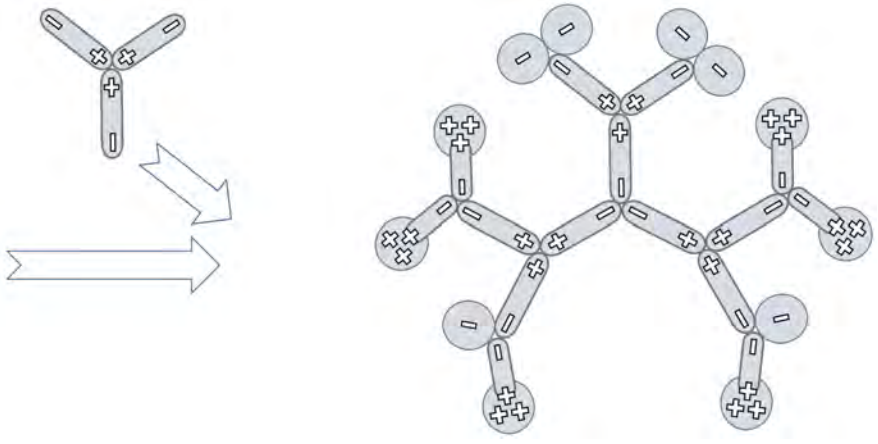


Figure 13. Reassembled atomic kernel for a helium-4 from Figure 12 by adding a TSL.

Consideration of the electrical charges: The net summation of the positive electrical charges is 6 × {2/3} (six up quarks times 2/3) and the net summation of the negative electrical charges is 6 × {1/3} (six down quarks times 1/3). The net balance of all electrical charges turns out two positive (6 × {2/3} − 6 × {1/3} = 6/3 = 2).

Consideration of the mass: The net summation of all six up and all six down quarks results a complete mass of four atomic units (six times a third from up quarks, plus six time a third from down quarks 6 × {1/3} + 6 × {1/3} = 12/3 = 4).

The three positive triple nodes and three negative triple nodes in the interior of the aggregate are equilibrated in charges.

The same mechanism is used for holding up and down quarks together, which is a re-interpretation for the strong forces.

As next step, atomic kernel of the helium isotope He-3 is considered, firstly by disassembling a hydrogen H-1 and a deuterium H-2 kernel respectively, as displayed in Figure 14.

By adding one triple space link TSL and reassembling them all, the result helium kernel He-3 is resulted in Figure 15.

Consideration of the electrical charges: The net summation of the positive electrical charges is 5 × {2/3} (five up quarks times 2/3) and the net summation of the negative electrical charges is 4 × {1/3} (four down quarks times 1/3). The net balance of all electrical charges turns out two positive (5 × {2/3} − 4 × {1/3} = 6/3 = 2).

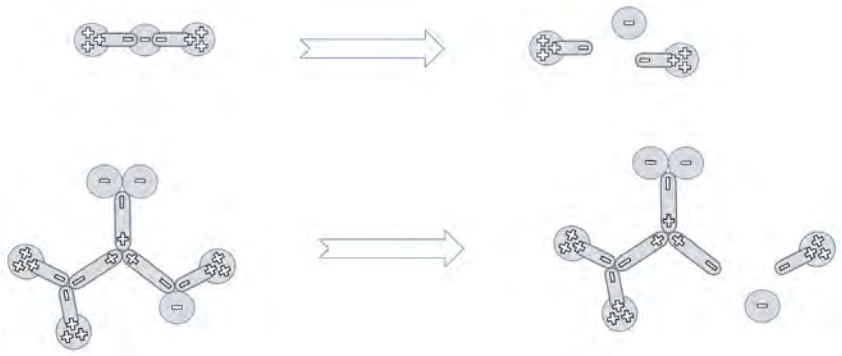


Figure 14. Disassembling hydrogen H-1 and deuterium H-2 kernel.

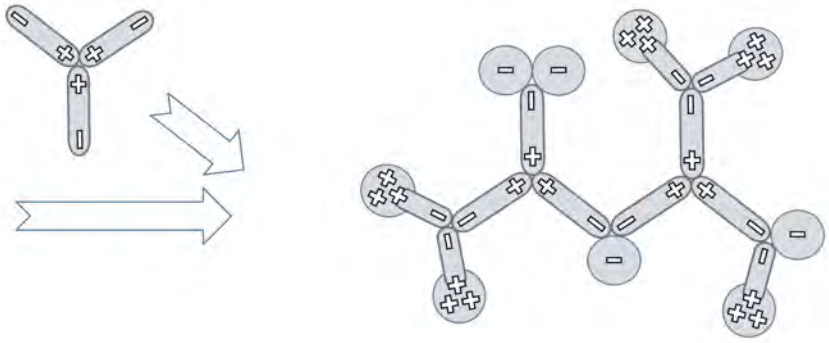


Figure 15. Reassembled atomic kernel for a helium-4 from Figure 12 by adding a TSL.

Consideration of the mass: The net summation of all five up and all four down-quarks results a complete mass of three atomic units (five times a third from up-quarks, plus four time a third from down-quarks $5 \times \{1/3\} + 4 \times \{1/3\} = 9/3 = 3$).

The two positive triple nodes and two negative triple nodes in the interior of the aggregate are equilibrated in charges.

8. Alternative Kernel Configurations for Helium He-4

YY model makes it possible that multiple consistent atomic kernel models (alternative configurations) for helium He-4 exist. One of the alternatives to Figure 13 is to reassemble the parts in Figure 12, by adding two triple space links, to Figure 16.

Another alternative is, also by adding two triple space links, to reassemble the parts from Figure 12-17.

All common things of models in Figure 13, Figure 16 and Figure 17 for helium atomic kernel He-4 are

- The numbers of up quarks (6) and the numbers of down quarks (6);
- The electrical charge balance (positive 2), and
- The mass summation units of all quarks (3).

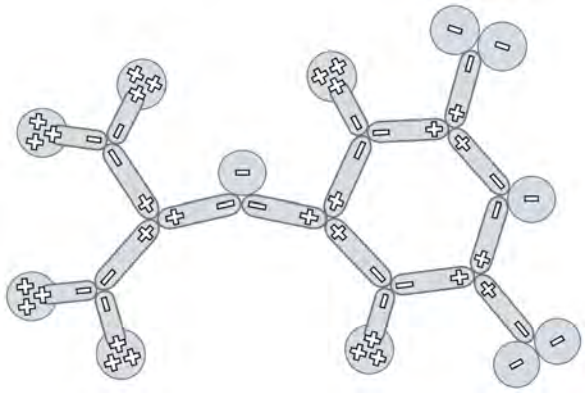


Figure 16. A Possible alternative atomic kernel model for helium He-4.

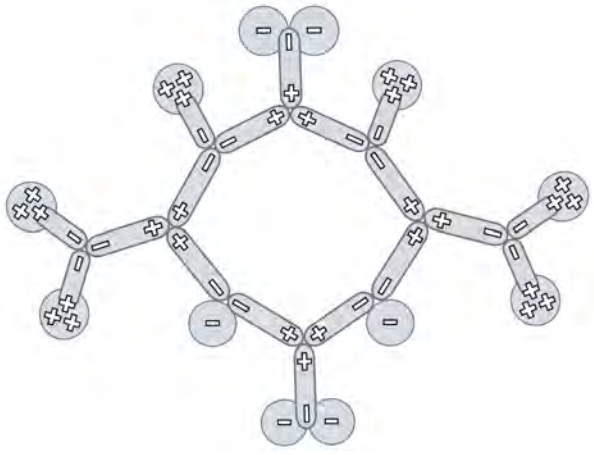


Figure 17. Another possible alternative atomic kernel model for Helium He-4.

The internal triple space links vary from alternatives: Figure 13 possesses three positive TSLs and their balancing negative nodes; Figure 16 and Figure 17 possess four positive TSLs and four balancing negative nodes respectively. Generally as a configuration rule:

- The number of positive TSLs is equal to the number of negative balancing nodes (**Internal Charge Balance ICB**).

ICB is important for the stability of aggregate structures, though the governing mechanism must be still investigated mathematically and physically.

Moreover, a TSL must possess a mass unit that is essential smaller in comparing to the mass of an up or down quark—these are in the order of a third atomic mass unit.

9. Process for Generating Electron and Positron

In this section, we discuss how the YY model understands the electron and positron, their generation processes using the prototype for the electron and splitting the quarks for the positron. The construction and composition of the electron and positron are found out here.

To generate an electron, the YY model formally considers a nuclear interaction of two neutrons with each other (dis- and reassembly). One proton and three "liberated" down quarks can be obtained, which together form a prototype for a negatively charged particle that decays successively into an electron and neutrino/antineutrino of different generations, as shown in Figure 18.

Essential aspects for generating electron are:

- An electron in the YY model is composed of deep-bound yin's, originated from three down quarks (likely a positive triple space link is composed of deep-bound yang's from three paring links);
- Kernel transformation model: 2n → p^+ + electron prototype → p^+ + e^- + neutrino, antineutrino.

The main difference between YY model and the well-known processes in standard model like muon decay ($\mu^- \to e^- + \bar{\nu}_e + \nu_u$) and beta decay ($n \to p^+ + e^- + \bar{\nu}_e$) (Ref. [11] and [13]) is the involved number of neutrons. Furthermore, the electron origination (prototype) and resulted properties (deep-bound three yin's) differ.

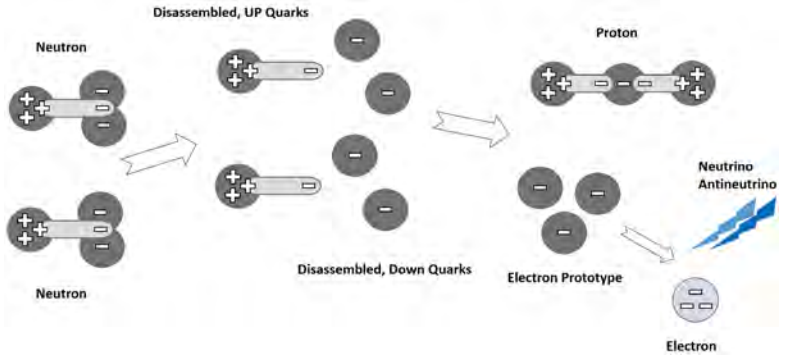


Figure 18. Generating process for electron.

Generating positron is interpreted by the "inverse beta decay" (Ref. [11]): $\bar{\nu}_e + p \rightarrow e^+ + n$, as displayed in Figure 19.

Obviously, the antineutrino strikes one up quark, shifted the mass (grey shadowed) from its positive charge pool to its negative pool becoming a down quark to combine the rest of the aggregate to a neutron, while the positive part of the concerned up quark becomes a positron.

Of course, the widely accepted standard particle model postulates that a free electron does not possess a measurable dimension and, that an electron is not dividable to smaller elementary particles.

From the perspective of YY model, electron as deep-bound yin's can still have a smaller dimension unit than our current physical detective ability. Furthermore, electron as "no dividable particle" is also a fact, with one difference:

The only way to reassemble an electron is its collision with a positron that leads to annihilation of the both particles and emission of gamma radiations. In the YY model, this collision will create a reassembled triple space link TSL (see next section) which is "given back" to the space and becomes a "waste" matter.

10. Electron-Positron Annihilation in YY Model, Outcome TSL and Y Particle

The YY model conserves the participating yin's and yang's in all transformation processes. The electron-positron annihilation described in standard model (Ref. [11] [14]) $e^+ + e^- \rightarrow 2\gamma$ will be extended in YY model by generating a TSL obeying this conservation, see Figure 20.

TSL is a prediction of the YY model. It can be imagined as ashes after the wood has burned out. Each TSL has a mass defined as a "TSL mass unit", which must be much smaller than the mass of electron and positron, but still not

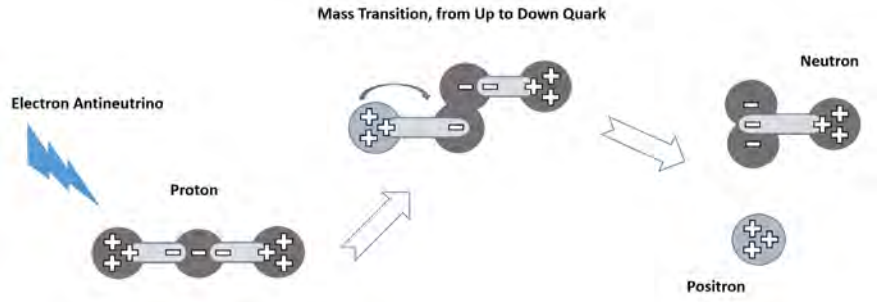


Figure 19. Generating process for positron.

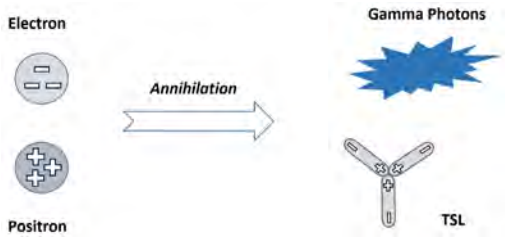


Figure 20. Electron-positron annihilation by generating gamma photons and a TSL.

negligible. It might be possible to determine it indirectly, using a more precise energy-matter balance between predictions and observations, also to estimate the absolute value of the TSL unit of mass.

Actually, electron-positron annihilation is part of the energy-matter life cycle and TSLs go back to their original form, which is given at the beginning of cosmological evolution. The third to last section will give a rough description of this life cycle. In addition, the TSL formation forms the basis for the antimatter systematics of the YY model (outside the scope of this paper).

We call TSL as **Y particle**, in order to motivate physicians to verify its existence.

11. Rules for Electrical Charge Units and Mass Units of Atomic Kernel

The early modelling approaches in this paper have showed, that the net electrical charges of an atomic kernel turn out from the summation of all building outer vertexes – all up and down quarks combinations. For clearness, some more terminologies are introduced here (see also Figure 21).

- **Neutronhead**: Triple end node containing two down quarks—possesses one negative electrical charge unit and 2/3 atomic mass unit;
- **Protonhead**: Triple end node containing one down quark—possesses one negative electrical charge unit and 1/3 atomic mass unit;
- **Protonid**: Single end node containing one up quark—possesses one positive electrical charge unit and 1/3 atomic massunit.

Additionally to the properties for electrical charges and mass units described above, there are also rules for balancing the numbers of protons and neutrons in an atomic kernel:

- Every protonhead needs two corresponding protonids for building a proton within an atomic kernel aggregate (number of protonheads = number of protons);
- Every neutronhead needs one corresponding protonid for building a neutron within an atomic kernel aggregate (number of neutronheads = number of neutrons).

Thus, the occurrences of protonids within an atomic kernel aggregate are not determining factor for calculating the numbers of protons and neutrons: They are only for satisfying the other requirements from the occurrences of protonheads and neutronheads.

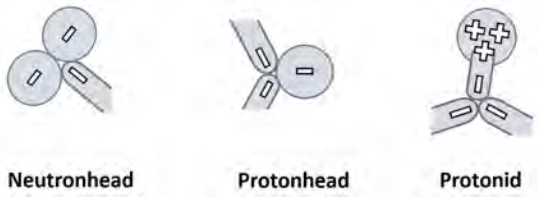


Figure 21. Neutronhead, protonhead and protonid.

This implies another asymmetry between up quark and down quark: The occurrences of down quarks (single or in pair) are essential for determining the protons and neutrons assembled in an atomic kernel.

Furthermore, a special rule has to be obeyed by the whole aggregate:

- The number of positive triple space links (TSLs) and the number of negative binding vertexes must balance in total (=> Internal Charge Balance ICB), as already discussed in the early Figure 9, Figure 11, Figure 13, Figure 16 and Figure 17.

In Figure 22 for the He-4 atomic kernel, colored texts annotate the internal charge balance (blue), respectively the external charge balance (red). "Internal" means pairing space links (PSLs) embedded within triple space links (TSLs). "External" means all surrounding protonheads, neutronheads and protonids.

The internal charge balance (ICB) guarantees the electrical neutrality of the binding element structure (PSLs, TSLs) so that the whole electrical charge balance of an atomic kernel just results from the summation of all surrounding protonheads, neutronheads and protonids.

All these rules give us the guidelines for assembling and reassembling new atomic kernels by disassembling the original atomic kernels, in a consistent way.

The atomic mass unit calculation for a kernel aggregate is simple and straightforward:

- The atomic mass units of an atomic kernel results from the summation of all atomic mass units of surrounding protonheads, neutronheads and protonids, divided by three;
- The TSL mass units results from the number of existing TSLs.

In Figure 23 for the He-4 atomic kernel, colored texts annotate atomic mass units of all protonheads, neutronheads and protonids (green), respectively the TSL mass units (blue).

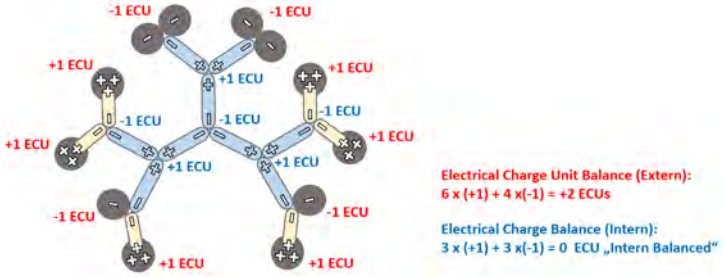


Figure 22. Electrical charge balances of helium kernel He-4.

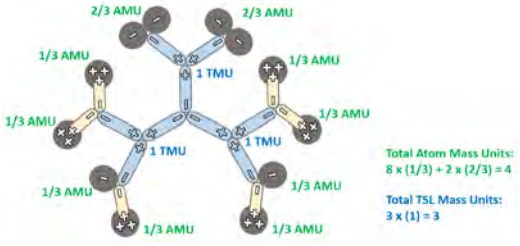


Figure 23. Atomic mass units and TSL mass units of helium kernel He-4.

12. Stellar Nuclear Fusion Process

This section describes some important parts of the stellar fusion reactions (Ref. [15] and [16])—divided into four sub-sections. The mechanisms of the YY model discussed in the early sections are applied here to understand the internal structural changes in the fusion reactions without considering details of the energy transitions.

12.1. Nuclear Fusion to Deuterium H-2

The beginning stage of stellar nuclear fusion produces a hydrogen isotope H-2 by merging two protons H-1 and generating a positron and neutrino:

$\text{H-1} + \text{H-1} \rightarrow \text{H-2} + \text{e}^{+} + \upsilon$.

Following the same construction procedure by disassembling the both H-1 and reassembling their elements to get H-2 get applied here. To ensure the model consistence by obeying the rules (last section), a new triple space link TSL is needed (**Figure 24**).

The complete transformation can be seen in two steps. In the first step, a proton is virtually transformed into a neutron by an inverse beta decay (**Figure 19**) and, in a further step, the virtual neutron and a second proton are merges into a deuterium (**Figure 9**) by adding a TSL.

The interpretation of the first step is equivalent to assuming the transition of an up quark to a down quark, a positron and an electron neutrino ($\text{u} \rightarrow \text{d} + \text{e}^{+} + \upsilon_{\text{e}}$), **Figure 25**.

However, what happens to the excessive positron? Moreover, where does TSL come from? Obviously, the both questions have something strongly to do with

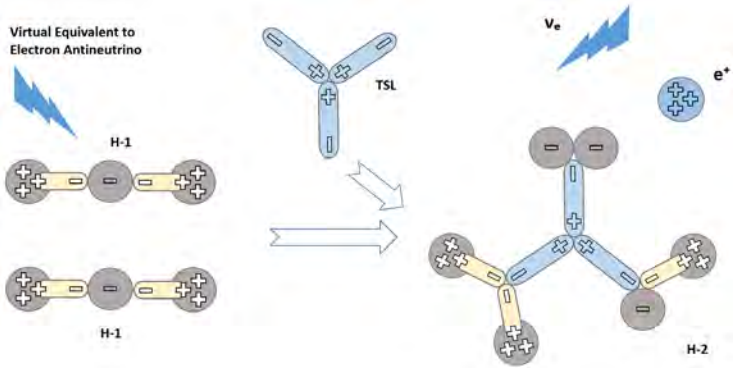


Figure 24. Nuclear fusion to H-2.

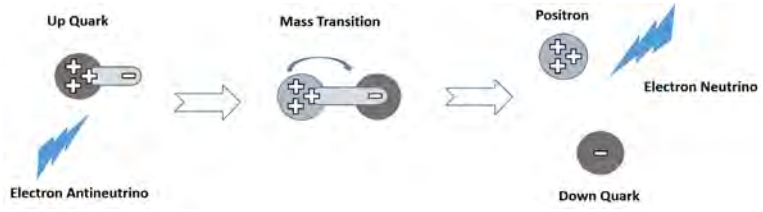


Figure 25. Transition of an up quark to down quark, positron and neutrino.

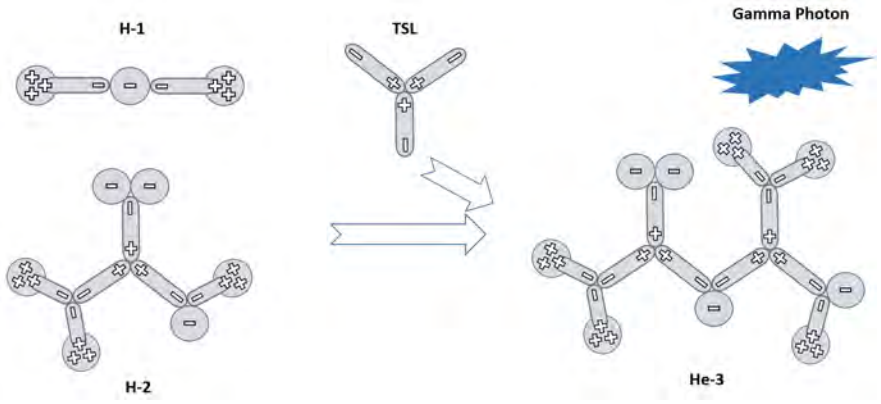


Figure 26. Nuclear fusion to He-3.

each other. This will be explained in the Section 12.4, "Mechanism for Producing TSLs in Stellar Nuclear Fusion".

12.2. Nuclear Fusion to Helium Isotope He-3

The next stage of stellar nuclear fusion produces a helium isotope He-3 by merging a proton H-1 and a hydrogen isotope H-2, emerging gamma photon: $\text{H-1}+\text{H-2}\rightarrow\text{He-3}+\gamma$ (Figure 26). Following the YY model, a new TSL is needed, accompanying the gamma emerging.

12.3. Nuclear Fusion to Helium He-4

In a further stage of stellar nuclear fusion, a stable helium He-4 and two protons H-1 will be produced by merging two helium isotopes He-3:

$\text{He-3}+\text{He-3}\rightarrow\text{He-4}+2\text{p}+12.9\ \text{MeV}$.

The structure of helium He-3 is already described in early section (Figure 15). The fusion process after YY model is described in Figure 27. To illustrate clearly, and for better identifying, parts are colored in aggregates before and after the fusion.

Be aware that this fusion stage releases an existing TSL, which can be re-consumed in the both of early fusion stages for H-2 and He-3.

12.4. Mechanism for Producing TSLs in Stellar Nuclear Fusion

Two stellar nuclear fusion stages consume TSLs, as already described in Section 12.1 and 12.2 (Figure 24 and Figure 26). This section explains the mechanism for producing TSLs: It comes from two sources:

- Released from the fusion stage to build He-4 (see Figure 27), and
- Generated from the beginning fusion stage to build H-2—this will be described in following.

The generation mechanism is composed of two nuclear transformation parts inside the fusion process $\text{H-1}+\text{H-1}\rightarrow\text{H-2}+\text{e}^{+}+\upsilon$:

✓ Virtual inverse beta decay (Figure 19): $\bar{\nu}_{\text{e}}+\text{p}\rightarrow\text{e}^{+}+\text{n}$

✓ Electron-proton annihilation (Figure 20): $\text{e}^{+}+\text{e}^{-}\rightarrow 2\gamma+\text{TSL}$

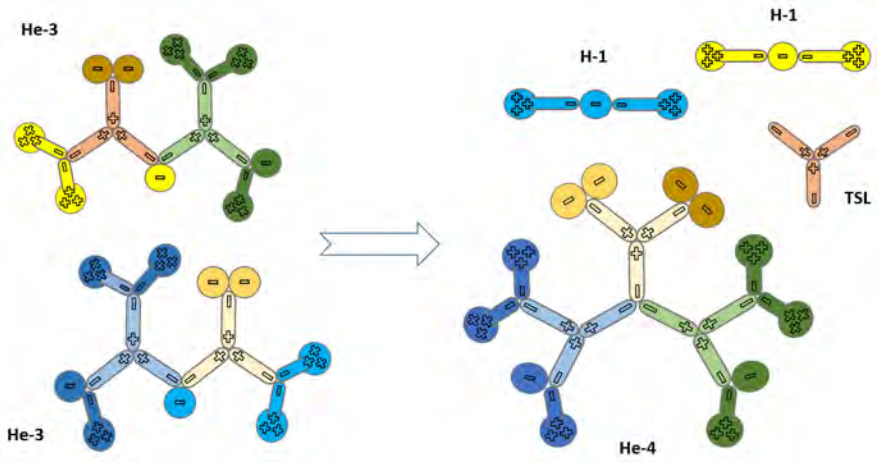


Figure 27. Nuclear fusion to He-4.

What happens is only possible to consider a pair of hydrogen atoms (two hydrogen kernels and two electrons) in a high-density state (Ref. [16]):

2H-1 + 2e$^-$, ionized & energized

$$\rightarrow \quad \bar{\nu}_e + \left(p^+ + e^-\right) + \left(p^+ + e^-\right) \quad \rightarrow \quad \left(e^+ + n + e^-\right) + \left(p^+ + e^-\right)$$

$$\rightarrow \quad \left(2\gamma + TSL + n\right) + \left(p^+ + e^-\right) \quad \rightarrow \quad 2\gamma + \left(TSL + n + p^+\right) + e^-$$

$$\rightarrow \quad 2\gamma + H\text{-}2 + e^-$$

With the considered group of two hydrogen atoms (already ionized), one up quark of the first H-1 is transformed to a neutron, producing one positron which is soon annihilated together with the first electron, producing gamma photons and a TSL. By consuming this TSL, the second H-1 merges with the transformed neutron to a deuterium kernel H-2 (Figure 9). The second electron balances this merged kernel electrically. The standard description ($H\text{-}1 + H\text{-}1 \rightarrow H\text{-}2 + e^+ + \upsilon$) gets here a more precise and detailed formulation.

The self-providing mechanism for TSL will also take effect for the second stage (Figure 26): $H\text{-}1 + H\text{-}2 \rightarrow He\text{-}3 + \gamma$, with the following expression concerning the participating hydrogen atom H-1:

$$\bar{\nu}_e + \left(p^+ + e^-\right) + \cdots \rightarrow \left(e^+ + n + e^-\right) + \cdots \rightarrow \left(2\gamma + TSL + n\right) + \cdots$$

Lastly, the third stage of fusion process for building stable helium He-4 releases a TSL, which can participate other fusion stages.

TSL emergency and its involvement in the nuclear fusion processes is new and important in comparison to the standard model. TSL is not easily detectable, thus out of focus of our experimental physics until now. It is a stable state of matter occurrence.

13. Essentials of a Nuclear Fission Process

A very essential thing happened in association with a nuclear fission process will be considered in this section: How a separation occurs at the fraction place, based on the YY model.

Consider the example nuclear fission of uranium atoms (Ref. [11]):

$$n + U\text{-}92/235 \rightarrow Ba\text{-}56/139 + Kr\text{-}36/95 + 2n + 200\ MeV$$

Without modelling the complete big atomic kernel of uranium 235 (that would massively extend the scope of this paper), the single place is interested which separates the uranium atom kernel into the both parts of Ba-139 and Kr-95.

The YY model will describe this place with a possible structure configuration before and after separation in **Figure 28**. Colored parts help to identify them in the structure of big kernel before and kernel pieces after the separation.

Not only up and down quarks are preserved, but the electrical charges before and after the separation. Additionally, two existing triple space links TSLs are released.

Allowedly, the model description above supposes, that there is a single place to separate the U-235 atom kernel to its fission products Ba-139 and Kr-95. It is possible the separation occurs cross multiple connected places – the final answers will turn out after modelling the complete uranium atoms. Many different valid structure configurations are possible, as the simple cases for tritium H-3 (**Figure 11**), and for helium He-4 (**Figure 13**, **Figure 16** and **Figure 17**) already illustrated.

Cold Fusion:

The YY model provides additional arguments that can help to understand whether cold fusion (Ref. [17]) can work.

A nuclear reaction without the consumption of TSL can be considered as "cold" because no electron-positron annihilation takes place and therefore no gamma radiation is produced. Transformation processes that even release TSLs can offer them for parallel running fusions, which in turn consume TSLs without producing them themselves.

Stellar fusion processes that require TSLs for binding, as already described for the formation of H-2 and He-3, must themselves "produce" TSLs and thus gamma photons.

Because the nuclear fission ($n + U\text{-}92/235 \rightarrow Ba\text{-}56/139 + Kr\text{-}36/95 + 2n + \cdots$)

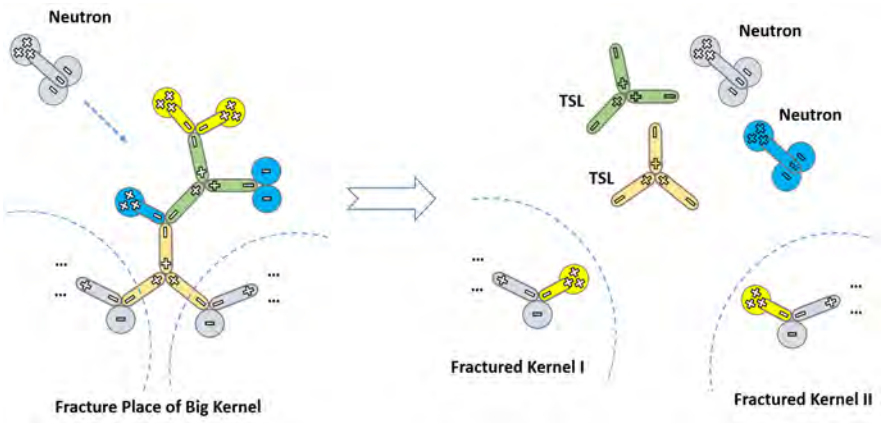


Figure 28. Fracturation of a big atomic kernel and its products.

also releases TSLs, it is possible that other big atomic kernels (not necessarily uranium) also do this so that these TSLs can be consumed by parallel running fusions without (or with few) gamma radiations.

Figure 29 is the YY model description for well-known nuclear fusion process for helium He-4 (Ref. [16]): $\text{H-2} + \text{H-3} \rightarrow \text{He-4} + 3.5\ \text{MeV} + (n + 14.1\ \text{MeV})$.

No TSL is involved in this process, neither consumed nor released. Gamma radiations ($e^- + e^+ \rightarrow \text{TSL} + 2\gamma$) do not occur. Thus, hydrogen fusion from H-2 and H-3 to helium He-4 becomes a good technical candidate for nuclear cold fusion, being interesting for human applications.

It should be mentioned, that the considerations here are only from the structure point of view. Energy balance aspects are not touched.

14. A Possible Scenario for Matter Evolution during Cosmological Inflation

If, as already mentioned, we go back to the origin of TSLs from the beginning of the cosmological evolution, we find some concrete construction processes that are closely related to the inflation phase (Ref. [18]). They are described in the following subsections as life cycle parts of a possible scenario.

14.1. Triple Space Link and Its Role for the Homogeneous Space

At an early stage during the inflation phase after the big bang, before the build-up of real matter, space should have a homogeneous structure caused by a high energy density. The YY model considers this homogeneous structure as composed of interconnected TSLs (called "Triple Space Linked State" TSLS). It is the initial state for further matter formation, Figure 30.

It is mainly a spatial structure based on electrostatic interactions of TSLs. The matter-antimatter asymmetry of the cosmos can already be founded here. At this stage of cosmological expansion, space is "cross-linked" by interconnected TSLs - The cosmos is filled with Y-particles.

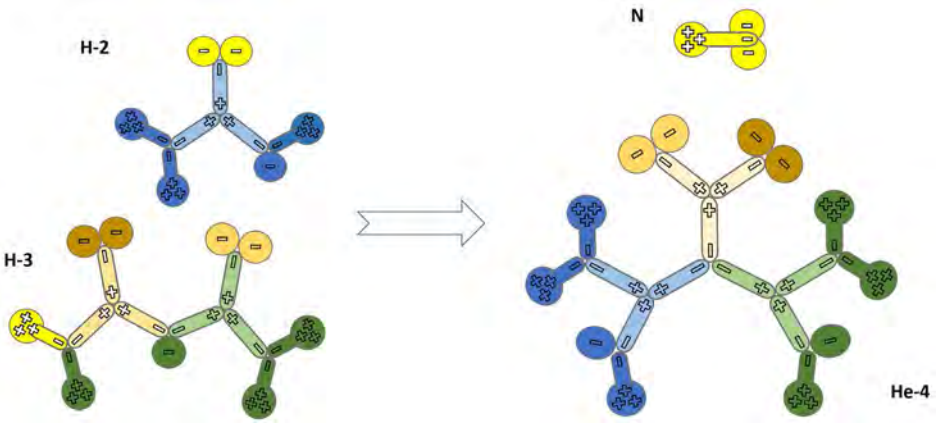


Figure 29. Fusion of deuterium and tritium to helium, accompanied by a neutron.

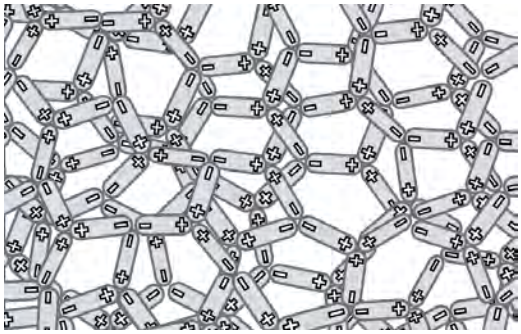


Figure 30. Triple space linked state.

14.2. Pre-Build Matter in Form of up and down Quarks from TSLs

In this stage, inter-linked TSLs will be released from each other. Up quarks and down quarks will be created from TSLs as manifestation of energy to matter: Every TSL itself, forced by energy, can be "transformed" into one up quark and two down quarks – cutting two yin's and fill all parts with matter, see Figure 31.

Because up and down quarks are dependent matter form, "pre-build" of matter is used here. At this stage, space is filled with up quarks, twice times numbers more with down quarks and rest TSLs, Figure 32.

14.3. Build Neutrons

The following stage builds neutrons, whereas the space is expanding and more transparency for particles and more freedom for pre-build matters given. One neutron is assembled by "tripling" one up and two down quarks, Figure 33.

Building neutrons can be regarded as one scenario. It is also possible that protons are built with existing up and down quarks parallel to neutrons.

14.4. Build Protons and Electrons

As already described in Figure 18 of Section 9, pairs of neutrons can reassemble themselves to protons and electrons. The universe filled with neutrons (Figure 33) will than descend soon into a universe filled partially with protons, electrons and rest neutrons (Figure 34):

Neutrinos and antineutrinos are emerged (not drawn in the Figure) by electron prototypes as mediator.

14.5. Pairing of Protons and Electrons to Hydrogen Atoms H-1

With expansion of space, free electrons are bound together with protons to form hydrogen atoms (Figure 35). Later hydrogen atoms are grouped to hydrogen molecules by chemical binding.

Two essential details still need to be explored: First, how cosmological inflation leads to a state "filled with" triple space links by taking into account how high energetic interactions take place while space expands at a very early stage of inflation. Secondly, how the energies manifest themselves in the subsequent expansion process to form pre-fabricated matter (up and down quarks).

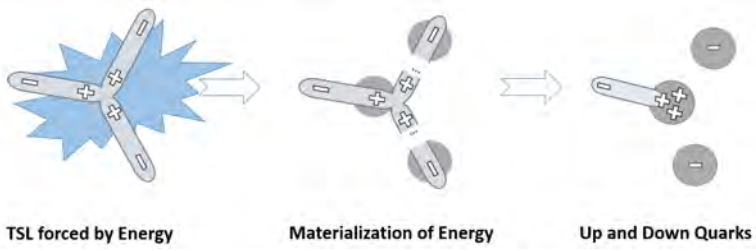


Figure 31. From one TSL to one up quark and two down quarks.

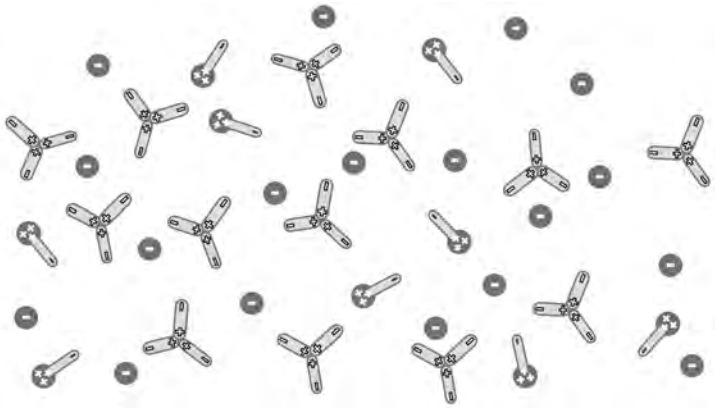


Figure 32. Universe filled with up quarks, down quarks and TSLs.

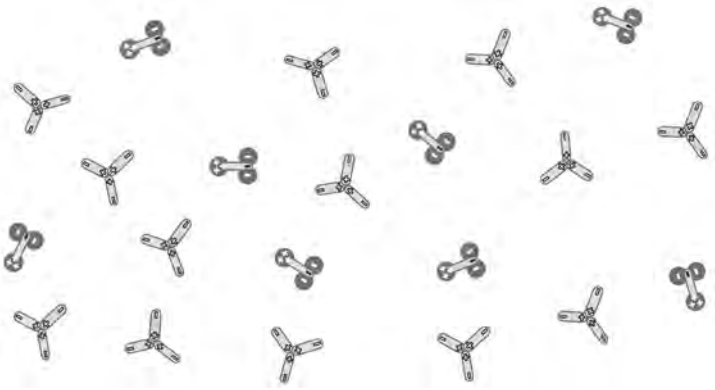


Figure 33. Universe filled with neutrons and TSLs.

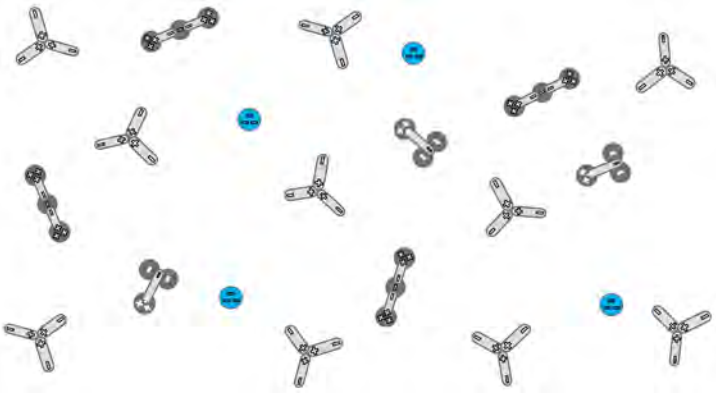


Figure 34. Universe filled with protons, electrons and neutrons, accompanied by TSLs.

Figure 35. Universe filled with hydrogen atoms and rest neutrons and TSLs.

15. Dark Matter Consideration

The dark matter postulated from cosmological research is still outside the experimental range of our detection technologies. TSL in the YY model would be a candidate, or it isclosely related to dark matter.

Triple space links TSLs, as a released product of Triple Space Linked State (TSLS, Section 14.1), form "low-state" or "reaction-lazy" stable matter. They are postulated to have a small mass unit and, more importantly, to be charge neutral. They are the initial structure for the formation of matter. High energetic concentrations force them to build up and down quarks or to combine with them for producing real matter. They are also the product of electron-positron annihilation, as described in Section 9, accompanied by gamma rays. As shown in previous sections, they are released during some nuclear fusion and fission processes (**Figure 27** and **Figure 28**). They can be consumed by other fusion processes (**Figure 24** and **Figure 26**).

Apparently, free TSLs are lost outside our current observation and detection technologies. Bounded TSLs inside of atomic nuclei are interpreted by existing theories of interactions. It is possible to model their occurrence and consumption during the nuclear transformation and annihilation processes. Their role in the construction of the nucleus and their number are known. It should be possible to find some experimental methods that could verify the physical existence of TSL (the Y-Particle) and its mass unit by balancing all measurable masses and energies.

16. Conclusions and Outlook

The YY model opens up a broad spectrum of theoretical and experimental research in atomic physics and molecular chemistry. On the basis of the detailed construction structure of the atomic nucleus and its variants, electrostatic and quantum field calculations could deliver more filigree results than is currently possible. For example, how is the three-dimensional structure of a heavy atomic nucleus contracted in some parts and stretched or folded in other parts by elec-

trostatic interactions between protonheads and neutronheads (both negatively charged) and protonides (positively charged).

So far, the governing forces for the PSL and TSL are hypotheses. A possible way to give them a basis is to combine them with QCD interpretation. Alternatively, similar methods from QCD can be used to clarify the necessary interactions in the YY model. The idea that the particle interactions are based on the cyclic groups Z2 (charges, materialized by the PSLs) and Z3 (colors, materialized by the TSLs) could be derived from this model for future analyses (Ref. [19]).

The distribution of electron clouds in orbitals around the atomic nucleus will also be important. A more detailed atomic model will reveal more details of the orbitals as well as molecular-level structures that have been hidden so far.

An isotope of an atom can have different nuclear structure configurations, as already shown for tritium H-3 (**Figure 11**) and for helium He-4 (**Figure 13**, **Figure 16** and **Figure 17**). The search for differences in physical and chemical properties can reveal interesting aspects and contribute to the research of radioactive decay processes.

The possible relationship of the YY model to the existing preon model (Ref. [20] and [21]) is an open question. In the preon model, the compositeness as substructure is suggested to go beyond the Standard Model.

Conflicts of Interest

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

References

[1] Gell-Mann, M. (1964) *Physics Letters*, **8**, 214. https://doi.org/10.1016/S0031-9163(64)92001-3

[2] Gaillard, M.K., Grannis, P.D. and Sciulli, F.J. (1999) *Reviews of Modern Physics*, **71**, S96. https://doi.org/10.1103/RevModPhys.71.S96

[3] Bortz, F. (2004) The Quark. The Rosen Publishing Group, Inc., New York.

[4] Epelbaum, E., Hammer, H.W. and Meißner, U.G. (2009) *Reviews of Modern Physics*, **81**, 1773-1825. https://doi.org/10.1103/RevModPhys.81.1773

[5] Machleidt, R., Holinde, K. and Elster, C. (1987) *Physics Reports*, **149**, 1-89. https://doi.org/10.1016/S0370-1573(87)80002-9

[6] Adamczyk, L., Adkins, J., Agakishiev, G., *et al.* (2015) *Nature*, **527**, 345-348. https://doi.org/10.1038/nature15724

[7] Burkert, V.D., Elouadrhiri, L. and Girod, F.X. (2018) *Nature*, **557**, 396-399. https://doi.org/10.1038/s41586-018-0060-z

[8] Chu, J. (2019) Study of Quark Speeds Finds a Solution for a 35-Year Physics Mystery. MIT News Office.

[9] Schmookler, B., Duer, M., Schmidt, A., *et al.* (2019) *Nature*, **566**, 354-358. https://doi.org/10.1038/s41586-019-0925-9

[10] Schmidt, A., Pybus, J., Weiss, R., *et al.* (2020) *Nature*, **578**, 540-544. https://doi.org/10.1038/s41586-020-2021-6

[11] Povh, B., Rith, K., Scholz, C., Zetsche, F. and Rodejohann, W. (2014) Particles and Nuclei. 9th Edition, Springer, Berlin. https://doi.org/10.1007/978-3-662-46321-5

[12] Ellwanger, U. (2012) From the Universe to the Elementary Particles. Springer, Berlin. https://doi.org/10.1007/978-3-642-24375-2

[13] Basdevant, J.-L., Rich, J. and Spiro, M. (2005) Fundamentals in Nuclear Physics—From Nuclear Structure to Cosmology. Springer, Berlin.

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

[15] Philips, A.C. (1999) The Physics of Stars. 2nd Edition, John Wiley & Sons, Hoboken.

[16] Morse, E. (2018) Nuclear Fusion. Springer, Berlin. https://doi.org/10.1007/978-3-319-98171-0

[17] Huizenga, J.R. (1992) Cold Fusion: The Scientific Fiasco of the Century. University of Rochester Press, Rochester.

[18] Allday, J. (2001) Quarks, Leptons and the Big Bang. 2nd Edition, CRC Press, Boca Raton. https://doi.org/10.1201/9780585312354

[19] Abbas, S.A. (2017) Group Theory in Particle, Nuclear, and Hydron Physics. CRC Press, Boca Raton. https://doi.org/10.1201/9781315371702

[20] Dugne, J.J., Fredriksson, S. and Hansson, J. (2002) *Europhysics Letters*, **57**, 188. https://doi.org/10.1209/epl/i2002-00337-8

[21] Fredriksson, S. (2004) Preon Prophecies by the Standard Model. *Proceedings of the Fourth Tegernsee International Conference on Particle Physics beyond the Standard Model*, Tegernsee, 9-14 June 2003, 211-223. https://doi.org/10.1007/978-3-642-18534-2_13

Journal of Modern Physics, 2020, 11, 1157-1169
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

An Improved Standard Model Comes with Explicit CPV and Productive of BAU

Chilong Lin

Department of Exhibition, National Museum of Natural Science, Taiwan
Email: lingo@mail.nmns.edu.tw, chilonglin@yahoo.com.tw

How to cite this paper: Lin, C.L. (2020) An Improved Standard Model Comes with Explicit CPV and Productive of BAU. *Journal of Modern Physics*, **11**, 1157-1169.
https://doi.org/10.4236/jmp.2020.118072

Received: June 30, 2020
Accepted: August 4, 2020
Published: August 7, 2020

Abstract

In this manuscript, we present an explicit way to describe the violation of CP symmetry in the standard model (SM) of electrweak interactions. In such a way, complex Cabbibo-Kobayashi-Maskawa (CKM) matrices are achieved which stand for the violation of CP symmetry. At the beginning, two necessary but not sufficient conditions for yielding a complex CKM matrix are stated as criteria. Then we found an interesting condition between the real and imaginary parts of a Hermitian 3×3 matrix which may provide extra relations among its parameters and reduce the number of them from eighteen down to five. In previous investigations, this can be done only down to nine. With another assumption among some of those parameters, the mass-matrix pattern is further simplified so as to be diagonalized analytically and in consequence four matrices which reveal S_N symmetries among or between quark generations are obtained. In some of such S_N-symmetric cases, the derived CKM matrices are complex which indicate that CP symmetry is violated accordingly. Taking the Jarlskog invariant as an estimate of the CPV strength, the value predicted by this model is orders stronger than the empirical value detected experimentally. However, that happens to fill partly the gap between the cosmologically observed amount of Baryon Asymmetry of the Universe (BAU) and that current Standard Model of particle physics predicts. It also proves the long suspected existence of BAU-productive eras in early universe if some fermions were indistinguishable, *i.e.*, S_N-symmetric, under circumstances with extremely high temperatures.

Keywords

CP Violation (CPV), Cabbibo-Kobayashi-Maskawa (CKM) Matrix, Baryon Asymmetry of the Universe (BAU), S_N Symmetry

1. Introduction

Since the first detection of CP violation (CPV) in 1964 [1], the issue of how CP

symmetry was violated, attracting our interest very much. However, for over fifty years, an explicit way to describe the violation of CP symmetry is still obscure. If we analyze the theoretical origin of CPV from its fundamental, considering only the quark sector here, the patterns of quark-mass matrices M are obviously the keys to ignite such a violation. Thus, we would like to analyze the CPV problem by starting from the most general pattern of an M matrix and then simplify it to a manageable level step by step.

In the standard model (SM) of electroweak interactions "Direct" CPV is allowed if a complex phase appears in the Cabbibo-Kobayashi-Maskawa (CKM) matrix describing quark mixings, or the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix describing lepton mixings. In SM such a complex phase can only be yielded by ingeniously arranging the Yukawa couplings between fermions and Higgs fields and then diagonalizing the Yukawa-coupling matrix directly. However, after decades such a satisfactory Yukawa-coupling matrix is still obscure.

Theoretically there is another way in SM to extract a complex phase from the vacuum expectation value (VEV) of its only Higgs doublet. However, such a phase can always be absorbed into redefinitions of quark fields in the SM. Thus, an extension of SM with one extra Higgs doublet was proposed in [2] which is usually referred to as the Two-Higgs-doublet model (2HDM). In this way, people expect the phases of two VEVs may unlikely be rotated away simultaneously if there were a nontrivial phase difference between them. However, such an extra Higgs doublet not only failed to solve the CPV problem, but also brought in an extra problem, the flavor-changing neutral current (FCNC) problem at tree level.

In fact, most of the derivations in this manuscript were originally proposed for solving the FCNC problem and CPV problem in the 2HDM [3]. However, we find they apply to the SM as well. As SM alone is enough to describe the violation of CP symmetry, why should one bother to deal with the extra Higgs doublet and problems it brings in? Thus, this manuscript is to be devoted to the theoretical ignition of CPV in the SM alone.

The Yukawa couplings of Q quarks in SM are usually given as

$$-L_Y = \bar{Q}_L Y^d \Phi d_R + \bar{Q}_L \epsilon Y^u \Phi^* u_R + h.c. \tag{1}$$

where Y^q are 3×3 Yukawa-coupling matrices for quark types $q = u, d$ and ϵ is the 2×2 anti-symmetric tensor. Q_L is left-handed quark doublets, and d_R and u_R are right-handed down- and up-type quark singlets in their weak eigenstates, respectively.

When the Higgs doublet Φ acquires a VEV, $\langle \Phi \rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$, Equation (1) yields mass terms for quarks with $M^q = Y^q v/\sqrt{2}$ the mass matrices for $q = u, d$. The physical states are obtained by diagonalizing Y^q with unitary matrices $U^q_{L,R}$ as $M^q_{diag.} = U^q_L M^q U^q_R = U^q_L Y^q U^q_R \left(v/\sqrt{2}\right)$, where L and R are their chirality.

As a result, the charged-current $W^{\pm}$ interactions couple to the physical u_L and d_L quarks with couplings given by

$$-L_{W^{\pm}} = \frac{g}{\sqrt{2}}\left(\bar{u}_L, \bar{c}_L, \bar{t}_L\right)\gamma^{\mu}W_{\mu}^{+}V_{CKM}\begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + h.c., \tag{2}$$

where

$$V_{CKM} = U_L^{u}\cdot U_L^{d\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{3}$$

is the CKM matrix describing quark mixings. Hereafter, we will neglect the sub-index L in quark fields q_L and the unitary matrices U_L if unnecessary.

If one demands a complex CKM matrix, following two conditions are necessary but not sufficient:

Condition **1**: At least one of either U^u or U^d is complex, *i.e.*, U^u and U^d musn't be both real.

Condition **2**: Even if both of them were complex, they musn't be the same, *i.e.*, $U^u \neq U^d$. Otherwise, V_{CKM} in Equation (3) will be a $1_{3\times 3}$ identity matrix which is obviously real.

As $V_{CKM} = U_L^{u}\cdot U_L^{d\dagger}$, obviously these two U matrices decide everything of V_{CKM} including if it were complex. As U^u and U^d are unitary matrices which diagonalize mass matrices M^u and M^d, respectively. It is obvious they are objects derived from M^u and M^d. Or, we may say the patterns of M^u and M^d are keys to ignite CPV in the standard model. Thus, we will start the investigation from a most general pattern of M matrix and then put in constraints to see what will happen to the CPV under various circumstances.

In section II, the most general pattern of M matrix is given as a start point of the investigation. If M were Hermitian, an interesting condition Equation (6) comes in if its real and complex parts can be diagonalized respectively by a same unitary matrix and the number of independent parameters in the M matrix is reduced from eighteen down to five. While in previous similar researches this can be done only down to nine with the same assumption.

Since only three quark masses are now given in a quark type, an M matrix with five unknowns is still too complicated to be diagonalized analytically. Thus, an assumption $A = A_1 = A_2 = A_3$, where A_1, A_2 and A_3 are diagonal elements of the most general M matrix to be given in section II, is employed to further simplify its pattern. Such an assumption gives us four analytically diagonalizable M_k matrices and correspondingly four complex U_k matrices, where k = 1 to 4 indicate to which case they correspond. These multiple U_k matrices enable us to satisfy both necessary conditions mentioned above for yielding a complex CKM matrix.

Even we have already a way to describe the violation of CP symmetry within the standard model. However, as to be demonstrated in Section III, the CPV de-

rived in such S_N-symmetric cases are orders stronger than the value current SM can provide and orders weaker than the one needed to account for the cosmologically observed Baryon Asymmetry of the Universe (BAU) [and references therein] [4]. The derived Jarlskog invariant $J \sim 0.171$, which is usually used to estimate the CPV strength [5], is about four orders stronger than the experimentally detected value $J = 3.00^{+0.15}_{-0.09} \times 10^{-5}$ [6] which corresponds to the BAU amount of order $\eta \sim 10^{-20}$ [7] and at least six orders weaker than the quantity $\eta = \left.\frac{N_B}{N_\gamma}\right|_{T=3\mathrm{K}} \sim 10^{-10}$ [6] [7] [8] estimated from the baryon/photon ratio observed in present universe.

At first glance such a CPV strength looks like a defect of the model. However, if we consider such S_N symmetries should exist only under circumstances of extremely high temperatures. It is not strange that CPV derived in this manuscript is different to the one detected in present experiments. Besides, it indicates the BAU we see today could be remnant left over in some early S_N-symmetric eras of our universe, at least part of them.

Conclusions and discussions are to be given in section IV.

2. Analytically Diagonalizable *M* Matrices and Their *U* Matrices

As mentioned in Equation (3), V_{CKM} is an inner product of two U matrices, one U^u from the up-type quarks and one U^{d+} from the down-type quarks, while they are respectively derived from the mass matrices M^u and M^d. Obviously these two M matrices are key factors to determine if V_{CKM} were complex. Based on this, the most orthodox way is to start from the most general pattern of M matrices and then diagonalize them to achieve corresponding U matrices.

In the SM with three fermion generations and one Higgs doublet, the mass matrix of any specific fermion type must has the general pattern as

$$M = \begin{pmatrix} A_1 + iD_1 & B_1 + iC_1 & B_2 + iC_2 \\ B_4 + iC_4 & A_2 + iD_2 & B_3 + iC_3 \\ B_5 + iC_5 & B_6 + iC_6 & A_3 + iD_3 \end{pmatrix} = \begin{pmatrix} A_1 & B_1 & B_2 \\ B_4 & A_2 & B_3 \\ B_5 & B_6 & A_3 \end{pmatrix} + i \begin{pmatrix} D_1 & C_1 & C_2 \\ C_4 & D_2 & C_3 \\ C_5 & C_6 & D_3 \end{pmatrix}, \quad (4)$$

which contains eighteen parameters and all A, B, C and D parameters are real.

Theoretically, the most orthodox way to derive the CKM matrix is to diagonalize one such matrix for up-type quarks and one for down-type quarks and then derive their corresponding U matrices. However, such a pattern is obviously too complicated to be diagonalized analytically. Not to mention the V_{CKM} thus derived contains thirty six parameters in total, eighteen from the up-type quarks and eighteen from the down-type quarks.

In many previous studies, physicists employed various *ad hoc* constraints to simplify the matrix pattern to a manageable level. For instance, the Fritzsch ansatz (FA) [9] [10] and its subsequent developments like Cheng-Sher ansatz (CSA) [11], Du-Xing ansatz (DXA) [12], combination of the Fritzsch and the

Du-Xing ansatz (DFXA and FDXA), combination of different assignments in the Du-Xing ansatz ($\tilde{\text{X}}$A), Non-mixing Top quark Ansatz (NTA) [and references therein] [13] and Fukuyama-Nishiura ansatz (FNA) [and references therein] [14] imposed several zeros in the M matrix as *ad hoc* constraints to reduce the number of parameters in it. The goal of all of them is just to simplify the M pattern to a manageable level. However, constraints so strong are in fact unnatural and unnecessary. As to be demonstrated in what follows, a very weak assumption that M matrices are Hermitian is already enough to yield a very simple M pattern and it includes almost previous ansatz as special cases in it.

It's interesting that action of "different assignments" in [13] and subsequent researches had in fact noticed the necessary condition 2. However, they did not state this condition explicitly like what is done here.

In SM, the Lagrangian as a whole is assumed to be Hermitian. If we assume the Yukawa couplings or equivalently the M matrices were also Hermitian, their patterns will be simplified remarkably. As shown in one of our previous investigations [15], such an assumption can reduce those eighteen parameters in Equation (4) down to only five.

In the first stage, the Hermitian assumption $M = M^{+}$ acquires that $D_j = 0$, $B_{j+3} = B_j$ and $C_{j+3} = -C_j$, for $j = 1,2,3$. Thus, Equation (4) becomes

$$\begin{aligned} M &= \begin{pmatrix} A_1 & B_1 + iC_1 & B_2 + iC_2 \\ B_1 - iC_1 & A_2 & B_3 + iC_3 \\ B_2 - iC_2 & B_3 - iC_3 & A_3 \end{pmatrix} \\ &= M_R + M_I = \begin{pmatrix} A_1 & B_1 & B_2 \\ B_1 & A_2 & B_3 \\ B_2 & B_3 & A_3 \end{pmatrix} + i \begin{pmatrix} 0 & C_1 & C_2 \\ -C_1 & 0 & C_3 \\ -C_2 & -C_3 & 0 \end{pmatrix} \end{aligned} \tag{5}$$

where M_R and M_I are the real and imaginary parts of M, respectively. Obviously, M_R and M_I are also respectively Hermitian. Thus, the number of parameters in M is reduced from eighteen down to nine at this stage.

If there were a unitary matrix U which diagonalize M_R and M_I simultaneously and respectively, it must also diagonalize the whole M matrix and leads to the following equation which was originally proposed in [16] and revised in [15],

$$M_R M_I^{\dagger} - M_I M_R^{\dagger} = 0 \tag{6}$$

Such an equation will provide us several extra relations among parameters and reduce the number of independent parameters further.

Substituting Equation (5) into Equation (6), we receive four equations

$$B_1 C_1 = -B_2 C_2 = B_3 C_3, \tag{7}$$

$$\left(A_1 - A_2\right) = \frac{B_3 C_2 + B_2 C_3}{C_1}, \tag{8}$$

$$\left(A_3 - A_1\right) = \frac{B_1 C_3 - B_3 C_1}{C_2}, \tag{9}$$

$$\left(A_2 - A_3\right) = -\frac{B_2 C_1 + B_1 C_2}{C_3}. \tag{10}$$

which will reduce the parameter number from nine further down to five.

However, five parameters are still too many to diagonalize the M matrix analytically. In order to simplify the pattern of M further, we will employ an extra assumption

$$A = A_1 = A_2 = A_3, \tag{11}$$

among the diagonal elements of M and further summarize Equations (7)-(10) in

$$B_1^2 = B_2^2 = B_3^2,\ C_1^2 = C_2^2 = C_3^2. \tag{12}$$

By examining all possible solutions of Equation (12) we found only four cases satisfy them. In what follows they will be studied respectively as:

Case 1: $B_1 = B_2 = B_3 = B$ and $C_1 = -C_2 = C_3 = C$

In this case,

$$M_1 = M_R + M_I = \begin{pmatrix} A & B & B \\ B & A & B \\ B & B & A \end{pmatrix} + i \begin{pmatrix} 0 & C & -C \\ -C & 0 & C \\ C & -C & 0 \end{pmatrix}, \tag{13}$$

which has the same pattern as what was achieved in [3] [17] with a S_3 permutation symmetry imposed among the Yukawa interactions of three quark generations. That symmetry was originally employed to solve the FCNC problem in a 2HDM. However, we find it applies to SM also.

Analytically, the M matrix is diagonalized and the mass eigenvalues are given as

$$M_{diag.} = \begin{pmatrix} A - B - \sqrt{3}C & 0 & 0 \\ 0 & A - B + \sqrt{3}C & 0 \\ 0 & 0 & A + 2B \end{pmatrix}, \tag{14}$$

and the U matrix which diagonalize Equation (13) is then given as

$$U_1 = \begin{pmatrix} \frac{-1 - i\sqrt{3}}{2\sqrt{3}} & \frac{-1 + i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{-1 + i\sqrt{3}}{2\sqrt{3}} & \frac{-1 - i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \tag{15}$$

where the sub-index k (k=1 to 4) of U_k indicates to which case it corresponds.

Case 2: $B_1 = B_2 = -B_3 = B$ and $C_1 = -C_2 = -C_3 = C$

In this case,

$$M_2 = M_R + M_I = \begin{pmatrix} A & B & B \\ B & A & -B \\ B & -B & A \end{pmatrix} + i \begin{pmatrix} 0 & C & -C \\ -C & 0 & -C \\ C & C & 0 \end{pmatrix}, \tag{16}$$

which reveals a residual S_2 symmetry between the second and third generations.

The mass eigenvalues are given as

$$M_{diag.} = \begin{pmatrix} A+B-\sqrt{3}C & 0 & 0 \\ 0 & A+B+\sqrt{3}C & 0 \\ 0 & 0 & A-2B \end{pmatrix}, \tag{17}$$

and corresponding U matrix is given as

$$U_2 = \begin{pmatrix} \frac{1-i\sqrt{3}}{2\sqrt{3}} & \frac{-1-i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1+i\sqrt{3}}{2\sqrt{3}} & \frac{-1+i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \tag{18}$$

Case 3: $B_1 = -B_2 = B_3 = B$ and $C_1 = C_2 = C_3 = C$

In this case,

$$M_3 = M_R + M_I = \begin{pmatrix} A & B & -B \\ B & A & B \\ -B & B & A \end{pmatrix} + i\begin{pmatrix} 0 & C & C \\ -C & 0 & C \\ -C & -C & 0 \end{pmatrix}, \tag{19}$$

which reveals a residual S_2 symmetry between the first and third generations.

The mass eigenvalues are now given as

$$M_{diag.} = \begin{pmatrix} A+B-\sqrt{3}C & 0 & 0 \\ 0 & A+B+\sqrt{3}C & 0 \\ 0 & 0 & A-2B \end{pmatrix}, \tag{20}$$

and corresponding U matrix is given as

$$U_3 = \begin{pmatrix} \frac{-1+i\sqrt{3}}{2\sqrt{3}} & \frac{1+i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{-1-i\sqrt{3}}{2\sqrt{3}} & \frac{1-i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \tag{21}$$

Case 4: $B_1 = -B_2 = -B_3 = -B$ and $C_1 = C_2 = -C_3 = -C$.

In this case,

$$M_4 = M_R + M_I = \begin{pmatrix} A & -B & B \\ -B & A & B \\ B & B & A \end{pmatrix} + i\begin{pmatrix} 0 & -C & -C \\ C & 0 & C \\ C & -C & 0 \end{pmatrix}, \tag{22}$$

which reveals a residual S_2 symmetry between the first and second generations.

The mass eigenvalues are given as

$$M_{diag.} = \begin{pmatrix} A+B-\sqrt{3}C & 0 & 0 \\ 0 & A+B+\sqrt{3}C & 0 \\ 0 & 0 & A-2B \end{pmatrix}, \tag{23}$$

and corresponding U matrix is then given as

$$U_4 = \begin{pmatrix} \frac{1-i\sqrt{3}}{2\sqrt{3}} & \frac{1+i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1+i\sqrt{3}}{2\sqrt{3}} & \frac{1-i\sqrt{3}}{2\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \tag{24}$$

In all four cases the U_k matrices are complex which satisfy the first condition stated in section I. If we assign different of them to up- and down-type quarks respectively, the second condition will also be satisfied. Even so, we are still not sure if V_{CKM} were complex since the inner product of two complex matrices can still be real, which is to be demonstrated in next section.

In fact, derivations presented above were originally proposed to solve the FCNC problem in the 2HDM by finding matrix pairs which can be diagonalized by a same U matrix simultaneously [18]. However, we found in some special cases the derived CKM matrices are complex which mean that CP symmetry is violated. Besides, the derivations apply not only to 2HDMs, but to the SM also. As SM is already enough to give CPV explicitly, why should we bother ourselves to deal with the FCNC problem in the 2HDM? Surely we can still apply them to the extensions of SM with one or even two extra Higgs doublets while the FCNC problem vanishes naturally at tree level. But that is a different subject to be discussed elsewhere.

Besides, it is noteworthy those S_N symmetries are not imposed *ad hoc* constraints. They are by-products brought in by the assumption among A parameters in Equation (11). Among them, the S_3-symmetric pattern in **case 1** is exactly the same as the one derived in our very early manuscripts [3] [17]. It solved the FCNC problem at tree level successfully but not the CPV problem. The problem it met was the breach of condition **2**, $U^u \neq U^d$, since we had only one S_3-symmetric U_1 matrix for both quark types at that time. While in this manuscript, extra S_2-symmetric U_2, U_3 and U_4 matrices provide us a key to ignite violation of CP symmetry from the theoretical end.

3. CKM Matrix and the BAU Problem

As mentioned in section I, if one expects to yield a CP-violating phase in the CKM matrix, two necessary conditions are to be satisfied. In last section, four complex U_k matrices were achieved with a Hermitian assumption of M and an assumption among A parameters. If we assign different U_k matrices to up- and down-type quarks respectively, both conditions are satisfied. In what follows, various assignments of U_k^u and U_k^d are examined and V_{CKM} they yield are presented in Table 1. Several of them are complex which indicate a theoretical origin of CPV is explicitly yielded.

The full expressions of matrices $1_{3\times3}, E, F, G$ and H in Table 1 are given as what follows

Table 1. Various assembles of CKM matrix.

V_{CKM}	$U_1^{d\dagger}$	$U_2^{d\dagger}$	$U_3^{d\dagger}$	$U_4^{d\dagger}$
U_1^u	$1_{3\times3}$	E	E^*	G
U_2^u	E^*	$1_{3\times3}$	H	F
U_3^u	E	H	$1_{3\times3}$	F^*
U_4^u	G	F	F^*	$1_{3\times3}$

$$1_{3\times3}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\ G=\begin{pmatrix}\frac{2}{3}&\frac{-1}{3}&\frac{2}{3}\\\frac{-1}{3}&\frac{2}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{2}{3}&\frac{-1}{3}\end{pmatrix},\ H=\begin{pmatrix}\frac{-1}{3}&\frac{2}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{-1}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{2}{3}&\frac{-1}{3}\end{pmatrix},$$

$$E=\begin{pmatrix}\frac{1-i\sqrt{3}}{3}&\frac{1}{3}&\frac{1+i\sqrt{3}}{3}\\\frac{1}{3}&\frac{1+i\sqrt{3}}{3}&\frac{1-i\sqrt{3}}{3}\\\frac{1+i\sqrt{3}}{3}&\frac{1-i\sqrt{3}}{3}&\frac{1}{3}\end{pmatrix}=\begin{pmatrix}\frac{2}{3}e^{-i\frac{\pi}{6}}&\frac{1}{3}&\frac{2}{3}e^{i\frac{\pi}{6}}\\\frac{1}{3}&\frac{2}{3}e^{i\frac{\pi}{6}}&\frac{2}{3}e^{-i\frac{\pi}{6}}\\\frac{2}{3}e^{i\frac{\pi}{6}}&\frac{2}{3}e^{-i\frac{\pi}{6}}&\frac{1}{3}\end{pmatrix},$$

$$F=\begin{pmatrix}\frac{1}{3}&\frac{1-i\sqrt{3}}{3}&\frac{1+i\sqrt{3}}{3}\\\frac{1+i\sqrt{3}}{3}&\frac{1}{3}&\frac{1-i\sqrt{3}}{3}\\\frac{1-i\sqrt{3}}{3}&\frac{1+i\sqrt{3}}{3}&\frac{1}{3}\end{pmatrix}=\begin{pmatrix}\frac{1}{3}&\frac{2}{3}e^{-i\frac{\pi}{6}}&\frac{2}{3}e^{i\frac{\pi}{6}}\\\frac{2}{3}e^{i\frac{\pi}{6}}&\frac{1}{3}&\frac{2}{3}e^{-i\frac{\pi}{6}}\\\frac{2}{3}e^{-i\frac{\pi}{6}}&\frac{2}{3}e^{i\frac{\pi}{6}}&\frac{1}{3}\end{pmatrix},\qquad(25)$$

The matrices $1_{3\times3}, G$ and H are purely real and obviously CP-conserving. While E, F and their complex conjugates are complex, which means they are CP-violating. It is interesting that as shown in the G and H cases, even if both conditions were satisfied, V_{CKM} is still completely real.

We now have already a way to describe the violation of CP symmetry from the theoretical end. But the CKM elements derived in Equation (25) are very different to experimentally detected corresponding values. Some of them are hundreds times the detected values, say both predicted $|V_{ub}|=2/3$ in E and F are about 184 times the presently detected value $3.82\pm0.24\times10^{-3}$ [6]. Besides, the CKM matrices derived here contain only numbers rather than any variable factors. That leaves us no space to improve the fitting between theoretical predictions and experimental detections further. This could be ascribed to the over-simplified matrix patterns caused by the $A=A_1=A_2=A_3$ assumption in Equation (11) or equivalently the revealed S_N symmetries. However, it may also hint that if we can diagonalize Equation (5) directly rather than imposing the assumption in Equation (11), we will achieve M patterns more close to the ones

observed now.

As we have already yield several complex V_{CKM} demonstrated in **Table 1**, it is rational for us to go one step further to estimate the CPV strength predicted by such a model. In usual, the strength of direct CPV is estimated by the dimensionless Jarlskog determinant [4] [5] [19] [20] which were given as

$$\begin{aligned}\Delta_{CP} &= v^{-12} ImDet\left[m_u m_u^{\dagger}, m_d m_d^{\dagger}\right] \\ &= Jv^{-12}\prod_{i<j}\left(\tilde{m}_{u,i}^2 - \tilde{m}_{u,j}^2\right)\prod_{i<j}\left(\tilde{m}_{d,i}^2 - \tilde{m}_{d,j}^2\right) \cong 10^{-19},\end{aligned} \tag{26}$$

where v is the Higgs vacuum expectation value and $\tilde{m}$ are particle masses.

The Jarlskog invariant "J" in Equation (26) is a phase-convention-independent measure of CPV defined by

$$\mathrm{Im}\left[V_{ij}V_{kl}V_{il}^{*}V_{kj}^{*}\right] = J\sum\nolimits_{m,n}\epsilon_{ikm}\epsilon_{jln}, \tag{27}$$

and a global fit of its value was given by the **Particle Data Group** as $J = 3.00^{+0.15}_{-0.09}\times 10^{-5}$ [6]. Such a value corresponds to a BAU of the order $\eta \sim 10^{-20}$ if one substitutes the detected fermion masses into Equation (26). It is obviously too small to account for the cosmologically observed

$$\eta = \left.\frac{N_B}{N_\gamma}\right|_{T=3\mathrm{K}} \sim 10^{-10}$$

where N_B is the number of baryons and N_γ is the number of photons.

Substituting the values of V_{cd}, V_{ts}, V_{cb}^{*} and V_{us}^{*} given in Equation (25) into (27), the J values for F and F^{*} are the same

$$J_{F,F^{*}} = \frac{16}{81}\sin\left(\frac{2\pi}{3}\right) \sim 0.171, \tag{28}$$

while those for E and E^{*} are zero since the phases in them are canceled. The J value given in Equation (28) is almost four orders higher than the one given by current SM. It hints that under circumstances with $S_2^{(c\leftrightarrow t)} + S_2^{(d\leftrightarrow s)}$, $S_2^{(u\leftrightarrow t)} + S_2^{(d\leftrightarrow s)}$, $S_2^{(u\leftrightarrow c)} + S_2^{(d\leftrightarrow b)}$ and $S_2^{(u\leftrightarrow c)} + S_2^{(s\leftrightarrow b)}$ symmetries, the hyper-indices in the brackets indicate between which two quarks the S_2 symmetry appears, the CPV strengths thus derived are orders stronger than what we see nowadays and thus more productive of BAU.

But, as constrained by Equation (27), the value of J is always smaller than 1. So a large J is always not enough to account for all the discrepancy between the cosmologically observed BAU and the one current SM predicts. One another potential source of a large Δ_{CP} is the mass-depending terms $\left(\tilde{m}_{u,i}^2 - \tilde{m}_{u,j}^2\right)$ and $\left(\tilde{m}_{d,i}^2 - \tilde{m}_{d,j}^2\right)$ in Equation (26). Unfortunately, the mass eigenvalues derived here give no helps on this subject since they are fixed to the detected quark masses. Obviously, finding mass eigenvalues running with some variables in such models will be very helpful in solving the BAU problem.

Besides, symmetries usually exist under circumstances with higher temperatures. As we do not see such S_N symmetries among fermion generations in our present universe, it is natural for us to consider they could have appeared in

some early eras if the Big-Bang cosmology were correct. Thus, the discrepancy between the S_N-symmetric CPV and that detected in present experiments is very natural.

In another way, we may imagine in some very early stages of the universe with extremely high temperature T there were S_3 symmetries in both quark types. As T decreased with the expansion, some fermions degenerated from others and the symmetries were broken down to a residual S_2-one. Probably the up-type quarks first and then the down-type quarks followed, but this is not necessary. During a succession of breaking down of S_N symmetries, for instance

$$S_3^u + S_3^d \to S_2^u + S_3^d \to S_2^u + S_2^d \to No^u + S_2^d \to No^u + No^d \tag{29}$$

where the super-indices u and d indicate up- and down-type quarks respectively. At least the $S_2^u + S_2^d$ era was proved to be CP-violating and very productive of BAU in this manuscript.

As we do not see any S_N symmetries in our present universe, obviously we are now in the latest stage, $No^u + No^d$, of Equation (29). That means a completely analytical diagonalization of M matrices is desired. We expect such a diagonalization will give us a V_{CKM} coincides the experimentally detected values better. Unfortunately, this is still unaccomplished for now and will be the next goal of our future investigations.

4. Conclusions and Discussions

Since the discovery of CPV in decays of neutral kaons, its theoretical origin is always a puzzle physicists' urgent to solve. Through the analysis on compositions of V_{CKM} two necessary but not sufficient conditions for yielding a complex V_{CKM} are stated and an explicit way to describe the violation of CP symmetry in SM is presented.

The way performed in this manuscript is not a new one to study the CP problem. It's in fact an old and orthodox way in SM, by diagonalizing the M matrices to obtain corresponding U matrices and then yields a complex V_{CKM}. In previous attempts we failed in this goal since most researchers ignored the second condition while in this manuscript that condition is satisfied since four such U_k matrices are achieved.

The key factor enables us to achieve a manageable M pattern and consequently a CP-violating V_{CKM} is Equation (6). Such a condition between the real and imaginary parts of a Hermitian matrix is always true as proved in [15] if they were diagonalized by a same U matrix. It correlates the elements in an M matrix and thus reduces the number of parameters from eighteen down to five. Furthermore, with an extra assumption among its parameters as given in Equation (11) four S_N-symmetric M_k patterns are achieved in Section II. Thus, the second condition is satisfied and several complex CKM patterns come in explicitly.

Though the magnitudes of Jarlskog invariants thus derived are about four orders that predicted by current SM. However, as we do not see any such S_N

symmetries in our present universe and they should exist only under circumstances with extremely high temperature. It's natural to consider their appearances in some very early stages of the universe. That hints our universe could have been very productive of BAU in such S_N-symmetric stages. While, as the S_N symmetries were broken down completely as the universe expanded, such a BAU-productive mechanism was narrowed down to the present status, $J = \left(3.00^{+0.15}_{-0.09}\right) \times 10^{-5}$. Obviously, such an improved SM not only describes how CP symmetry was violated from the theoretical end, but also explains why the amount of BAU observed today is so huge while the currently detected CPV is so tiny.

This work is not a new model. Instead, it is just an improvement of SM in the sector of fermions' Yukawa interactions. Though it does not solve the CPV and BAU problems completely, it still improves our understanding of the origin of CPV and exhibits a way to account for partly the cosmologically observed BAU which cannot be accounted for by current SM. Such an interesting aspect could be a good stepping stone for coming researchers on these topics.

Conflicts of Interest

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

References

[1] Christenson, J.H., Cronin, J.W., Fitch, V.L. and Turlay, R. (1964) *Physical Review Letters*, **13**, 138-140. https://doi.org/10.1103/PhysRevLett.13.138

[2] Lee, T.D. (1973) *Physical Review D*, **8**, 1226-1239. https://doi.org/10.1103/PhysRevD.8.1226

[3] Lin, C.L., Lee, C.E. and Yang, Y.W. (1988) *Chinese Journal of Physics*, **26**, 180-193.

[4] Tranberg, A., Hernandez, A., Konstandin, T. and Schmidt, M. (2010) *Physics Letters B*, **690**, 207-212. https://doi.org/10.1016/j.physletb.2010.05.030

[5] Jarlskog, C. (1985) *Physical Review Letters*, **55**, 1039-1042. https://doi.org/10.1103/PhysRevLett.55.1039

[6] Zyla, P.A., *et al.* (2020) *Progress of Theoretical and Experimental Physics*, **2020**, 083C01.

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

[8] Ade, P.A.R., *et al.* (2016) *Astronomy & Astrophysics*, **594**, A13.

[9] Fritzsch, H. (1978) *Physics Letters B*, **73**, 317-322. https://doi.org/10.1016/0370-2693(78)90524-5

[10] Fritzsch, H. (1979) *Nuclear Physics B*, **155**, 189-207. https://doi.org/10.1016/0550-3213(79)90362-6

[11] Cheng, T.P. and Sher, M. (1987) *Physical Review D*, **35**, 3484-3491. https://doi.org/10.1103/PhysRevD.35.3484

[12] Du, D. and Xing, Z.Z. (1993) *Physical Review D*, **48**, 2349-2352. https://doi.org/10.1103/PhysRevD.48.2349

[13] Carcamo, A., Martinez, R. and Rodriguez, J. (2007) *The European Physical Journal C*, **50**, 935-948. https://doi.org/10.1140/epjc/s10052-007-0264-0

[14] Matsuda, K., Nishiura, H. and Fukuyama, T. (2000) *Physical Review D*, **61**, Article ID: 053001. https://doi.org/10.1103/PhysRevD.61.053001

[15] Lin, C.L. (2019) *Journal of Modern Physics*, **10**, 35-42. https://doi.org/10.4236/jmp.2019.101004

[16] Branco, G.C., Buras, A.J. and Gerard, J.M. (1985) *Nuclear Physics B*, **259**, 306-330. https://doi.org/10.1016/0550-3213(85)90638-8

[17] Lee, C.E., Lin, C.L. and Yang, Y.W. (1990) *Physical Review D*, **42**, 2355-2358. https://doi.org/10.1103/PhysRevD.42.2355

[18] Lin, C.L. (2013) Type-IV Two-Higgs-Doublet Models and Their CP-Violation. (Unpublished) arXiv:1308.6039
Lin, C.L. (2014) A Different Origin of CP-Violation in Two-Higgs-Doublet Models. (Unpublished) arXiv:1403.3482

[19] Farrar, G.R. and Shaposhnikov, M.E. (1993) *Physical Review Letters*, **70**, 2833-2836. https://doi.org/10.1103/PhysRevLett.70.2833

[20] Farrar, G.R. and Shaposhnikov, M.E. (1994) *Physical Review D*, **50**, 774-818. https://doi.org/10.1103/PhysRevD.50.774

Journal of Modern Physics, 2020, 11, 1170-1179
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Magnetic Charge Theory. Part 3: Neutron Structure and Dark Matter Source

Keith G. Lyon

Haymarket, VA, USA
Email: klhaymarket@aol.com

How to cite this paper: Lyon, K.G. (2020) Magnetic Charge Theory. Part 3: Neutron Structure and Dark Matter Source. *Journal of Modern Physics*, **11**, 1170-1179.
https://doi.org/10.4236/jmp.2020.118073

Received: July 10, 2020
Accepted: August 7, 2020
Published: August 10, 2020

Abstract

The concept of magnetic charge is further developed to explain the electron/proton magnetic bonding that forms the neutron. The derivation leads to a minimum range for the Coulomb force of 2.35 fm that explains the lack of the Coulomb force in the nucleus. Further investigation into the nature of gravity leads to the possibility that dark matter is a byproduct of stars.

Keywords

Neutron, Dark Matter, Gravity, Coulomb Force, Magnetic Monopole, Magnetic Charge

1. Introduction

The concept of a magnetic charge that produces a magnetic field proportional to its velocity was introduced to explain dark matter [1] [2], a high interest topic in physics with no explanation. Magnetic charge is distinct from a magnetic monopole as shown in Figure 1. If dark matter is given a positive magnetic charge then a transverse wave through the dark matter produces the electromagnetic properties of light, thereby equating dark matter with the ether. Assuming this ether is an ideal gas in an adiabatically expanding universe produced the result that the speed of light was related to the ether density. This result not only explained how the universe's expansion would be interpreted as accelerating when it is decelerating but also showed that spacetime was related to the ether density. When considering the fundamental ground state vibration of a particle with both electric charge and a negative magnetic charge, the particle produced two new types of waves in the ether that had the properties to explain the transmission of the Coulomb force (vortex photons) and gravity (gaussian photons).

Magnetic Charge Theory also concluded that mass is a calculated property of

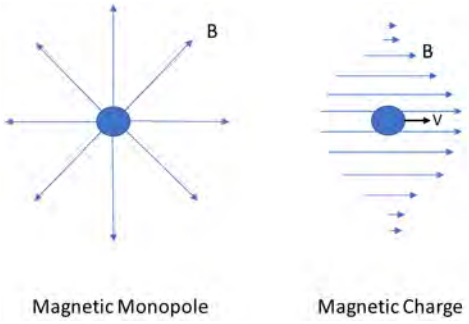


Figure 1. Magnetic monopole/magnetic charge distinction [2].

a particle that is proportional to the square of the magnetic charge [2]. A more profound result was obtained by calculating the total electromagnetic energy of a magnetic charge, vibrating in its ground state, and finding that this energy was equivalent to its rest mass energy. In other words, the rest energy of the magnetic particle was all electromagnetic energy.

This paper will further develop the result that total rest energy of particle is electromagnetic, and investigate if this result can be extended to understand the proton, neutron and electron. Specifically can these particles' rest energies be explained as purely electromagnetic? This investigation develops insight into electric and magnetic fields from charged particles that eliminate their field singularities at short distances. In addition, this work provides additional insight into gravity that was developed in a previous paper [2].

2. Electromagnetic Singularities

To calculate the total electromagnetic energy of a charged particle, the singularities that arise when calculating the electric field energy from an electric charge must be addressed. The energy in an electric field, E_E, is given by [3]:

$$E_E = \frac{1}{8\pi}\int \boldsymbol{E}\cdot\boldsymbol{D}\mathrm{d}V \tag{1}$$

where $\boldsymbol{E}$ is the electric field and $\boldsymbol{D}$ is the Displacement equal to $\varepsilon\boldsymbol{E}$, where ε is the dielectric constant (equal to one in this application).

The integral diverges when inserting the electric field from an electrically charged particle, q_e/r^2, and integrating from zero to infinity. Similarly, a moving electric charge creates a magnetic field [3]

$$\boldsymbol{B} = \frac{q_e}{c}\frac{\boldsymbol{v}\times\boldsymbol{x}}{|\boldsymbol{x}|^3} \tag{2}$$

The energy in a magnetic field is given by [3]:

$$E_B = \frac{1}{8\pi}\int \boldsymbol{B}\cdot\boldsymbol{H}\mathrm{d}V \tag{3}$$

This integral diverges for a moving, electrically charged particle when it is evaluated from zero to infinity. These divergences are not new and indicate that the singularities are inconsistent with reality, especially the coexistence of protons in

the nucleus.

Magnetic charge provides an explanation that eliminates these singularities. Electrically charged particles exchange vortex photons to exert the Coulomb force [2]. One possible explanation considers that two protons approaching each other will exchange vortex photons until the vortex photons are larger than the separation and unable to transmit the Coulomb force. Another possible explanation that could prevent attraction is that the vortex photon formation is disturbed when the two charges are too close. In either case there is a range where the electric field is reduced from the inverse square relationship. While the exact form of the electric field is unclear, a simple approximation is that the field remains inversely proportional to the separation squared outside a cutoff range, r_{co}, and zero inside this range. This leads to the electric field energy given by:

$$E_E = \frac{q_e^2}{8\pi}\int_{r_{co}}^{\infty}\frac{1}{r^4}\mathrm{d}V = \frac{q_e^2}{2r_{co}} \tag{4}$$

For an electron, setting E_E to its rest mass energy, mc^2, yields a cutoff range of 1.4 fm, however, this does not consider the electromagnetic field energies due to the vibrating ground state. It does, though, indicate that the cutoff range must be greater than 1.4 fm since vibrating charges contribute additional energy above just static electric field energy. For the proton, a similar calculation yields a cutoff range three orders magnitude smaller, but logically the cutoff range should not be dependent upon magnetic charge (or mass).

The magnetic field singularity from a moving electric charge has a different explanation. The electric field is related to the flux density of vortex photons travelling at the speed of light. The magnetic field is related to the ether density traveling at a velocity up to the speed of light. Magnetic fields approaching a singularity would mean that the ether is superluminal, but the velocity of the ether should be limited to the velocity of light. The magnetic field from a magnetic charge was previously assumed to be given by the following equation [2]:

$$\boldsymbol{B} = q_m \boldsymbol{v} \mathrm{e}^{-(\rho+|z|)/\sigma} \tag{5}$$

where the velocity is in the z direction, and ρ and z are cylindrical polar coordinates centered at the magnetic charge, q_m. Thus, if q_{met} is the magnetic charge of a particle of the ether, then the largest magnetic field, $\boldsymbol{B}_{\max}$, would be given by:

$$\boldsymbol{B}_{\max} = q_{met} c. \tag{6}$$

Thus, the magnetic field from a moving charge could follow Equation (2) until the field equals $\boldsymbol{B}_{\max}$. Inside the range where this occurs, the field would decrease to zero when the range goes to zero. The value of $\boldsymbol{B}_{\max}$ is currently unknown since q_{met} is unknown. Another interesting result is that this volume of reduced magnetic field near a moving electric charge is larger as the velocity of the charge particle increases.

In the case of lower velocities, there is another limit. With the electric field approximation being zero inside r_{co}, then its time derivative would also be zero inside r_{co} and hence, from Ampere's Law, the magnetic field is also zero. This

leads to the same approximation as the electric field, namely, the magnetic field is zero inside r_{co}. Thus, at lower velocities, inserting Equation (2) into Equation (3) and integrating from r_{co} to infinity, the magnetic field energy, E_M, is given by:

$$E_M = \frac{q_e^2 v^2}{3 r_{co} c^2} \tag{7}$$

3. Rest Mass Energy

Previous work showed that the rest mass energy of a magnetically charged particle was equal to the total electromagnetic energy that results from the particle vibrating in its ground state [2]. The analysis that follows will extend this observation by assuming that all energy is electromagnetic in nature and use this assumption to calculate r_{co} in Equation (4).

There are several components of the electromagnetic fields of a particle. First, there is the obvious electric filed due to its electric charge. The remaining fields are due to the vibration of its ground state, *i.e.*, the vibration of the electric and magnetic charges of the particle. The electric charge produces a circular magnetic field about the velocity direction, which is a maximum at maximum speed and zero at minimum speed. This changing magnetic field also yields an electric field in the velocity direction that is zero at maximum velocity and a maximum at zero velocity. Conservation of energy implies that the energy due to these vibrations is a constant, thus the energy is constantly transforming between magnetic and electric fields.

Similarly, the vibrating magnetic charge is creating magnetic and electric fields in its motion. The resulting magnetic field is in the velocity direction with a maximum at maximum speed and zero at zero speed. The changing magnetic field produces a changing electric field, circular about the velocity direction, that is maximum at zero speed and zero at maximum speed.

In both cases, the magnetic and electric fields are synchronized with the magnetic fields being a maximum at maximum speed and the electric fields being a maximum at zero speed. In addition, the magnetic fields produced by the magnetic charge are orthogonal to the magnetic fields produced by the electric charge; similarly, the electric fields are also orthogonal (but not orthogonal to the static electric field).

The total energy of the particle is the sum of the electric field energy and the magnetic field energy, which are both integrals of the fields squared over all space. The transformation of the vibrational energy between electric and magnetic fields complicates the integral for the electric field energy since the static field is not orthogonal to the vibrating electric fields. This would tend to imply that the total energy of the particle is fluctuating in defiance to the conservation of energy. There are two approaches to understand why this is not the case. First, if the vibrational electric fields are significantly contained inside the cutoff range, then the complication can be ignored, *i.e.*, $\sigma \ll r_{co}$. Second, the static electric field is a flux of vortex photons that are created by the ground state vibration of

the particle and would not exist simultaneously with the vibrational fields near the particle. Either of these explanations provide an answer consistent with energy conservation.

As a result of these arguments, the total energy of a particle, E_T, can be expressed as:

$$E_T = E_E + E_{ee} + E_{me} + E_{em} + E_{mm} \tag{8}$$

where:

E_E = static electric field energy (Equation (4)),

E_{ee} = electric field energy from electric charge motion,

E_{me} = magnetic field energy from electric charge motion,

E_{em} = electric field energy from magnetic charge motion, and

E_{mm} = magnetic field energy from magnetic charge motion.

E_{ee}, E_{me}, E_{em}, and E_{mm} are functions of time whose sum is a constant. The total energy can be calculated at any time in the cycle and, if the time is chosen at maximum speed or zero speed, two of these terms will equal zero.

4. Neutron, Proton and Electron Rest Energy Analysis

The neutron decays into a proton, an electron, and an antineutrino plus some energy, so the notion that the proton and electron are bound somehow is not a stretch of the imagination. A proton in motion produces a magnetic field according to Equation (5) inside r_{co}, and inside r_{co} there would effectively not be an electric field from electric charges; so, any binding would need to be a function of the magnetic field. An electron traveling at the same velocity would have a similar situation, but its magnetic field would be lower by the ratio of their magnetic charges. However, the proton's magnetic field is reduced at a distance and there exists a surface about the proton where its magnetic field would equal that of the electron traveling at the same velocity. If the electron were slightly displaced from this surface it would be forced back to this surface (as briefly discussed in reference [1]), hence, the electron's motion would be a sort of harmonic motion about this equilibrium surface created by the proton. Interestingly, this equilibrium surface is independent of velocity, hence, the ground state vibration of the proton would create a constant equilibrium surface where the electron would be bound as the proton's velocity changes.

While the collocation of the proton and electron negates the effect of electric charges, the total fields from the neutron are from the magnetic charges. The magnetic field of the neutron, $\boldsymbol{B}_n$, is given by Equation (5) as:

$$\boldsymbol{B}_n = \left(q_{me} + q_{mp}\right)\boldsymbol{v}\mathrm{e}^{-(\rho+|z|)/\sigma} \tag{9}$$

where q_{me} and q_{mp} are the magnetic charges of the electron and proton. The total field energy oscillates between a pure magnetic field at maximum velocity [2], c, and pure electric field at zero velocity. The total field energy can be calculated at any time with the easiest calculation occurring at maximum velocity. Inserting Equation (9) into Equation (3) yields the following as the total field energy of the

neutron, E_{Tn}:

$$E_{Tn} = \frac{1}{16}\sigma^3\left(q_{me} + q_{mp}\right)^2 c^2 = m_n c^2 \tag{10}$$

Expressions can be developed using a similar process for the electron and proton. In these cases, the electric charge contributes to the total energy, so the process is more involved. In both cases the static electric field energy is given by Equation (4).

The fields from their magnetic charges are also straight forward. Just as in the neutron calculation, the total field energy calculation can be performed at any time in the oscillation and the point of maximum velocity is again chosen. Inserting Equation (5) for the electron and proton into Equation (3) yields the magnetic field energy for the magnetic charge motion for the electron, E_{mme}, and proton, E_{mmp} at maximum velocity as:

$$E_{mme} = \frac{1}{16}\sigma^3 q_{me}^2 c^2 \tag{11}$$

$$E_{mmp} = \frac{1}{16}\sigma^3 q_{mp}^2 c^2 \tag{12}$$

The energy term relating to the magnetic field energy from the motion of the electric charge is determined by Equation (7), thus the magnetic field energy for the electron, E_{mee}, and the proton, E_{mep}, at maximum velocity are:

$$E_{mee} = \frac{q_{ee}^2}{3r_{co}} \tag{13}$$

$$E_{mep} = \frac{q_{ep}^2}{3r_{co}} \tag{14}$$

where q_{ee} and q_{ep} are the electric charges for the electron and proton.

Combining Equations (4) and (11)-(14) into Equation (8) produces the total energy for the electron, E_{Te}, and for the proton, E_{Tp}:

$$E_{Te} = \frac{q_{ee}^2}{2r_{co}} + \frac{q_{ee}^2}{3r_{co}} + \frac{1}{16}\sigma^3 q_{me}^2 c^2 = m_e c^2 \tag{15}$$

$$E_{Tp} = \frac{q_{ep}^2}{2r_{co}} + \frac{q_{ep}^2}{3r_{co}} + \frac{1}{16}\sigma^3 q_{mp}^2 c^2 = m_p c^2. \tag{16}$$

Equations (10), (15), and (16) form three equations with four unknowns that result in the following:

$$\frac{1}{r_{co}} = \frac{3c^2}{10q_e^2 m_n}\left[4m_e m_p - \left(m_n - m_p - m_e\right)^2\right] = \frac{1}{2.35\ \text{fm}} \tag{17}$$

$$\sigma^{1.5} q_{me} = -4\left[m_e - \frac{m_e m_p}{m_n} + \frac{\left(m_n - m_p - m_e\right)^2}{4m_n}\right]^{1/2} = -4.97\times10^{-15}\,\text{g}^{1/2} \tag{18}$$

$$\sigma^{1.5} q_{mp} = -4\left[m_p - \frac{m_e m_p}{m_n} + \frac{\left(m_n - m_p - m_e\right)^2}{4m_n}\right]^{1/2} = -5.17\times10^{-12}\,\text{g}^{1/2}. \tag{19}$$

These three equations were evaluated with the following: electron mass of $9.1093837015128\ (28) \times 10^{-28}$ g [4], proton mass of $1.67262192369\ (51) \times 10^{-24}$ g [5], neutron mass of $1.67492749804\ (95) \times 10^{-24}$ g [6], and electron charge of $-4.80320451\ (10) \times 10^{-10}$ esu [4].

The result that $r_{co} = 2.35$ fm from Equation (17) is consistent with the atom; being slightly greater than the size of the nucleus, it explains the lack of the Coulomb force in the nucleus, *i.e.*, the protons do not repel each other, and r_{co} is also much smaller than the Bohr radius so that the understanding of the electron cloud defined by quantum mechanics is unperturbed.

The magnetic charge results provide more insight into the proton and electron. The proton's magnetic charge is about 1040 times that of the electron from Equations (18 and 19). Further examination of their energy partition provides the results in **Table 1**, which illustrates that the proton's energy is primarily a result of its magnetic charge while the electron's energy is primarily a result of its electric charge.

The values in **Table 1** were calculated with respect to their rest energies, mc^2, and the solutions from Equations (17)-(19). The contribution from the static electric field energy is given by Equation (4). The contribution from the vibrating electric charge is given by Equations (13) and (14). The total electric charge contribution is the sum of these two contributions. The total vibrating magnetic charge contributions are from Equations (11) and (12).

The magnetic charge ratio between the proton and electron implies that their separation is about 7σ. While 7σ would seem like a large separation, σ was previously estimated to be less than 6.2×10^{-19} cm [2], hence 7σ would be less than 4×10^{-18} cm. This small separation compared to nuclear size and r_{co} would seem to justify the approximation that the proton and electron were collocated in the neutron.

The magnetic charge solutions and the σ estimate, lead to the lower bound for the proton and electron magnetic charges of $1.55 \times 10^{+16}$ sc·sec·cm^{-3} and $1.49 \times 10^{+13}$ sc·sec·cm^{-3}, respectively. This means that the peak magnetic fields that occur for the proton and electron in their ground state vibrations are greater than $4.65 \times 10^{+26}$ gauss and $4.48 \times 10^{+23}$ gauss. These extremely large fields indicate that the bonding to form the neutron is extraordinarily strong while the spatial extent of these fields is extremely limited due to their exponential nature.

Maybe the most important insight from this calculation is the neutron decay. The neutron decays into a proton, an electron, and an antineutrino, however,

Table 1. Percent energy contributions to basic particles' rest energy.

Energy Source	Electron	Proton	Neutron
Static electric charge	59.90%	0.032%	0
Vibrating electric charge	39.93%	0.022%	0
Total electric charge	99.83%	0.054%	0
Total vibrating magnetic charge	0.17%	99.946%	100%

this calculation does not require a third particle to bind the proton and electron. The electron is dominated by the proton's field, but the electron still has a ground state vibration that can perturb its motion with respect to the equilibrium surface. The electron's deviation from the equilibrium surface is driven back to the surface everywhere except when the motion takes it outside the surface at the extreme furthest distance from the proton that is normal to its vibration direction. At this extreme, any motion away from the proton (normal to the velocity) would create a cyclical motion of the electron about the proton before it exits the vicinity of the proton. This cyclical motion has the basic properties needed to produce a vortex photon, which may indicate that the antineutrino is a high-energy vortex photon. Since the vortex photon has a strong electric field impulse, this would certainly be consistent with the antineutrino creating an ionization trail in a cloud chamber.

5. Gravity Insights

In the previous paper [2], gaussian photons, g-photons, were shown to produce properties that resembled gravity. The g-photons are impulses of magnetic field in the velocity direction that radiate from the ground state vibrations of a particle. When these g-photons interact with normal matter with their negative magnetic charge, they produce an attractive force, just like gravity. By the same token, if there are particles with a positive magnetic charge, they would be repelled. As a result, in the interaction with the ether by g-photons, the ether would not be attracted and, thus, cannot explain the higher ether densities implied by the stronger gravitational field. In addition, if the ether density is larger in the galaxy center, it is not clear that this would produce a larger gravitational field.

If gravity does not attract the ether, then the obvious solution is that the ether is being created or released from stars. If this is the case, then a star should have a magnetic monopole moment.

A higher ether density at galaxy center can explain the stronger gravitational force, but not as a direct result of the additional mass of the ether. The explanation lies in the variation of space-time that results from the higher ether densities. The higher ether density will speed time up [1] and thus the particles will vibrate faster and thus produce a higher g-photon flux. The higher flux produces a larger gravitation pull and explains the observations that have led to the dark matter hypothesis.

6. Summary

The introduction of magnetic charge and the assumption that a particle's rest energy is all electromagnetic, provides a rational explanation for the magnetic binding of a proton and electron to produce the neutron. In addition, the theory results in a cutoff range for the Coulomb force of 2.35 fm, the revelation that the proton's energy is largely due to its magnetic charge, while the electron's energy

is largely due to its electric charge, and the possibility that the antineutrino is a high-energy vortex photon. All of these results are consistent with known physics and are key to understanding the proton's lack of repulsion in the nucleus.

Additional inspection of the interactions of g-photons implies the possibility that dark matter/ether is being generated or released from stars and would predict stars as having an overall magnetic monopole moment.

Magnetic Charge Theory and these results continue to explain some of the most important questions in physics, including:

- Explaining the photon, its electromagnetic fields, and why the magnetic fields do not form circuits as usually seen in other situations [2].
- Explaining that the variations in space-time are due to variations in the ether density [1].
- Explaining that the speed of light is dependent on the ether density, while explaining that the measurement of the local speed of light will always produce the same result when measured locally [1].
- Explaining that the transmission of the Coulomb force is by a flux of vortex photons emitted from the ground state vibration of matter. The repulsion/attraction forces are explained by the rotation direction of the v-photons [2].
- Explaining that the Coulomb force disappears within 2.35 fm of an electric charge and, thus, explaining why protons do not repel each other in the nucleus.
- Explaining that gravity is a flux of gaussian photons from a particle's ground-state vibration and that the g-photons possess a magnetic field impulse that interacts with a particle's magnetic charge [2]. The theory suggests that the ether is produced by stars and that stars have a magnetic monopole moment.
- Explaining how the universe's expansion is seen as accelerating (due to the assumption of a constant speed of light) when in fact the expansion is decelerating [1].
- Explaining that a particle's rest energy is purely electromagnetic, and mass is a calculated property proportional to the square of its magnetic charge.
- Explaining the neutron as a magnetically bound state of an electron and proton, and suggests that antineutrinos are high-energy v-photons.

Conflicts of Interest

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

References

[1] Lyon, K.G. (2016) *Journal of Modern Physics*, **7**, 920-927. https://doi.org/10.4236/jmp.2016.79084

[2] Lyon, K.G. (2018) *Journal of Modern Physics*, **9**, 1381-1396.

https://doi.org/10.4236/jmp.2018.97083

[3] Jackson, J.D. (1962) Classical Electrodynamics. John Wiley & Sons, Inc., New York.

[4] Wikipedia (2020) Electron.

[5] Wikipedia (2020) Proton.

[6] Wikipedia (2020) Neutron.

Journal of Modern Physics, 2020, 11, 1180-1192
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Empirical Equation for the Gravitational Constant with a Reasonable Temperature

Tomofumi Miyashita

Miyashita Clinic, Osaka, Japan
Email: tom_miya@plala.or.jp

How to cite this paper: Miyashita, T. (2020) Empirical Equation for the Gravitational Constant with a Reasonable Temperature. *Journal of Modern Physics*, **11**, 1180-1192.
https://doi.org/10.4236/jmp.2020.118074

Received: July 14, 2020
Accepted: August 9, 2020
Published: August 12, 2020

Abstract

Ted Jacobson discovered that gravity was related to thermodynamics. However, the calculated temperature using the Boltzmann area entropy is still not reasonable. We searched and discovered an empirical equation for the gravitational constant with a reasonable temperature. The calculated value was 3.20 K, which is similar to the temperature of the cosmic microwave background of 2.73 K. Then, we examined Yasuo Katayama's theory. For this purpose, we introduced the modified Wagner's equation, which is compatible with Jarzynski equality. Finally, using Ted Jacobson's theory, we proposed our theory of gravity with the Gibbs volume entropy.

Keywords

Gravitational Constant, Wagner's Equation, Jarzynski Equality, Ted Jacobson's Theory

1. Introduction

Ted Jacobson attempted to relate gravity and thermodynamics [1]. Subsequently, the theory of entropic gravity [2] was studied by Erik Verlinde. However, the calculated temperature values obtained are still not reasonable.

There are two kinds of entropy. The first is the Boltzmann area entropy, and the second is the Gibbs volume entropy (S_{Gibbs}). The theory of entropic gravity is based on the Boltzmann area entropy. The temperature calculated using this approach was very small and not reasonable. According to Jarzynski, the Gibbs volume entropy (S_{Gibbs}) is much stronger than the Boltzmann area entropy [3]. Therefore, instead of using the Boltzmann area entropy, we tried to use the Gibbs volume entropy to explain our empirical equation.

In the area of solid-state ionics, the open circuit voltage is governed by Wagn-

er's equation. However, we noticed that during ion hopping, the chemical potential of ions should increase, and the electrical potential energy of ions should decrease. Both the values of the increase in the chemical potential and the decrease in the electrical potential energy are equal to the ionic activation energy. We noticed that the ionic activation energy can be defined as the solvation Hamiltonian mean force (ϕ) [4]. The solvation Hamiltonian mean force (ϕ) in the electrolytes can be defined from the separation of the Boltzmann distribution. Furthermore, the solvation Hamiltonian mean force (ϕ) can be measured experimentally.

In our theory, the energy of the rest mass slightly decreases after pair production. At the same time, the Gibbs volume entropy (S_{Gibbs}) should slightly increase. Both the values of the increase and the decrease are equal to the solvation Hamiltonian mean force (ϕ). However, the ratio between the solvation Hamiltonian mean force (ϕ) and the energy of the rest mass are too small and cannot be observed. We assumed that the solvation Hamiltonian mean force (ϕ) should increase with increasing distances. Then, the rest mass should increase with increasing distances. The Gibbs volume entropy (S_{Gibbs}) should increase with increasing distances.

The rest of the paper is organized as follows: in Section 2, we present our empirical equations and our conclusions. In Section 3, we attempt to explain our empirical equations using the modified Wagner's equation. In Section 4, using Ted Jacobson's method, we showed the obtained conclusions for our theory of gravity.

The schematic view of our project is shown in **Figure 1**.

The central concept from our experimental results is related to Wagner's equation in electrochemistry.

2. Empirical Equation for the Gravitational Constant

2.1. Our Empirical Equation

Our empirical equation is (quoted from Wikipedia https://en.wikipedia.org/wiki/Proton)

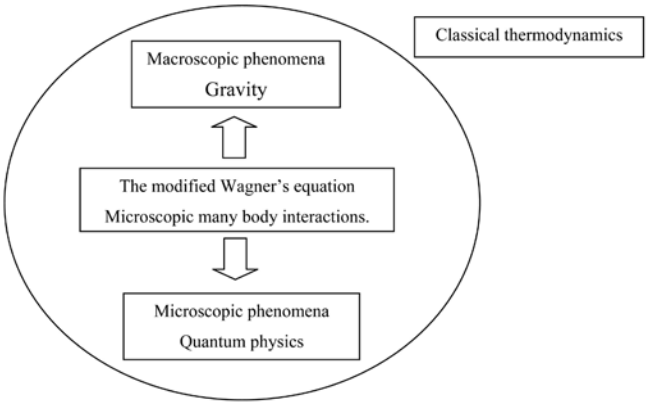


Figure 1. Schematic view of our project.

$$\frac{Gm_p}{r_p} \times 1\,\text{kg} = 3kT \tag{1}$$

G: gravitational constant, 6.6743×10^{-11} ($m^3 \cdot kg^{-1} \cdot s^{-2}$).
m_p: the rest mass of a proton, 1.67×10^{-27} (kg).
r_p: charge radius, 8.41×10^{-16} (m).
k: Boltzmann constant, 1.38×10^{-23} (J/K).
T: temperature (K).
1 kg: the standard mass (kg).

The temperature (T) calculated using this formula was 3.20 K, which is similar to the temperature of the cosmic microwave background of 2.73 K. The standard mass (1 kg) must be used due to the definition of the gravitational constant. This problem was explained in Section 3.1.

2.2. Our Conclusions from Our Empirical Equation

Equations (2)-(4) are our conclusions.

$$\phi = 3kT\,. \tag{2}$$

$$m_p c^2 = h\nu - 3kT \tag{3}$$

$$\frac{\mathrm{d}S_{\text{Gibbs}}}{\mathrm{d}r} T = \frac{3nkT}{r} \tag{4}$$

where h and ν are the Planck constant and the frequency, respectively. n is the number of protons, and r is the distance from a large mass ($n \times m_p$).

3. Explanation for Our Empirical Equation

3.1. Consideration of the Law of Universal Gravitation

This section is strongly influenced by Yasuo Katayama [5].
http://home.catv.ne.jp/dd/pub/pot/pot.html#0.

He thought that, according to Newton, the energy can be obtained infinitely from the gravitational potential energy, but according to Einstein, there should be an upper limit (mc^2). Thus, the rest mass should decrease with decreasing the gravitational potential. Equation (16) was quoted from Yasuo Katayama.

The gravitational potential ($\Phi(r)$) is given by

$$\Phi(r) = -\frac{GM}{r}\,, \tag{5}$$

where M and r are the mass and the distance from the mass, respectively. Then, the work (W) is given by

$$W = m \times \left(\Phi(r_2) - \Phi(r_1)\right) = m \times \left(\frac{GM}{r_1} - \frac{GM}{r_2}\right) \tag{6}$$

where m is a sufficiently small mass. r_1 and r_2 ($>r_1$) are the distances from mass M. Therefore,

$$W = \int_{r_1}^{r_2} m \times \frac{\mathrm{d}\Phi(r)}{\mathrm{d}r}\mathrm{d}r \tag{7}$$

Then,

$$\frac{\mathrm{d}W}{\mathrm{d}r} = mg \tag{8}$$

$$g = \frac{\mathrm{d}\Phi(r)}{\mathrm{d}r} \tag{9}$$

Next, from the mass-energy equivalence and the Law of the Conservation of Energy, we obtain

$$m_2 c^2 = m_1 c^2 + W \tag{10}$$

Therefore,

$$m_2 - m_1 = \frac{W}{c^2}. \tag{11}$$

m should increase with the increasing of r.

$$\frac{\mathrm{d}m}{\mathrm{d}r} = \frac{1}{c^2}\frac{\mathrm{d}W}{\mathrm{d}r}. \tag{12}$$

From Equations (5) and (9), we obtain

$$\frac{\mathrm{d}m}{\mathrm{d}r} = \frac{1}{c^2} mg\ . \tag{13}$$

Therefore, from Equation (6) and Equation (10), we find that

$$\frac{\mathrm{d}m}{\mathrm{d}r} = \frac{1}{c^2} m \frac{\mathrm{d}\Phi(r)}{\mathrm{d}r}. \tag{14}$$

Then,

$$\mathrm{d}m = \frac{1}{c^2} m \mathrm{d}\Phi(r). \tag{15}$$

Then, we assumed that m should be the minimum mass (m_{p1}) when r is the minimum distance (r_{p1}).

$$\ln\left(\frac{m_{p2}}{m_{p1}}\right) = \frac{1}{c^2}\left(\Phi\left(r_{p2}\right) - \Phi\left(r_{p1}\right)\right) \tag{16}$$

Thus,

$$\ln\left(\frac{m_{p2}}{m_{p1}}\right) = \frac{G}{c^2}\left(\frac{M_1}{r_{p1}} - \frac{M_2}{r_{p2}}\right) \tag{17}$$

When M is n (the number of protons) × m_p,

$$M_1 = n \times m_{p1} \tag{18}$$

$$M_2 = n \times m_{p2} \tag{19}$$

From Equations (17)-(19),

$$\ln\left(\frac{m_{p2}}{m_{p1}}\right) = \frac{nG}{c^2} \times \left(\frac{m_{p1}}{r_{p1}} - \frac{m_{p2}}{r_{p2}}\right) \tag{20}$$

when r_{p2} is infinitely large.

$$\ln\left(\frac{m_{p2}}{m_{p1}}\right)=\frac{nG}{c^2}\times\frac{m_{p1}}{r_{p1}} \tag{21}$$

Therefore,

$$\frac{c^2}{n}\ln\left(\frac{m_{p2}}{m_{p1}}\right)=G\times\frac{m_{p1}}{r_{p1}} \tag{22}$$

When M_1 is 1 kg, n is 6.02×10^{26} ($=1000\times10^{23}$). The value of n is calculated from Avogadro's number.

$$1\,\text{kg}=n\times m_{p1} \tag{23}$$

thus forming Equation (22) and Equation (23),

$$m_{p1}c^2\times\ln\left(\frac{m_{p2}}{m_{p1}}\right)=1\,\text{kg}\times G\times\frac{m_{p1}}{r_{p1}} \tag{24}$$

Then, n disappeared. This does not mean that n is unimportant. As shown in Equation (20), n is clearly meaningful. We tried to explain the influence of 1 proton on the gravitational potential, and then, n disappeared. When Newton defined 10 kg as the standard mass, the value of G should change. However, Equation (24) should not change. The real meaning of Equation (24) was discussed in the later section.

3.2. Consideration from Wagner's Equation

This section is the most important and is quoted from our own experimental results [4]. In this section, we tried to explain the concept of the solvation Hamiltonian mean force from the perspective of thermodynamics in relation to hopping conduction.

A solid oxide fuel cell (SOFC) directly converts the chemical energy of a fuel gas, such as hydrogen or methane, to electrical energy. A solid oxide film is used as the electrolyte, where the main carriers are oxygen ions. As electrolytes, YSZ and SDC (samaria-doped ceria) are well known. The open circuit voltage (OCV) is calculated from Wagner's equation.

$$J_{O_2}=-\frac{RT}{16F^2L}\int_{\ln\left(pO_2^{\text{anode}}\right)}^{\ln\left(pO_2^{\text{cathode}}\right)}\frac{\sigma_{\text{el}}\sigma_{\text{ion}}}{\sigma_{\text{el}}+\sigma_{\text{ion}}}\,\text{d}\ln\left(pO_2\right) \tag{25}$$

where J_{O_2} and pO_2 are the O_2 flux and the O_2 partial pressure, respectively; pO_2^{cathode} and pO_2^{anode} are the O_2 partial pressures at the cathode and the anode, respectively; R, T, and F are the gas constant, the absolute temperature, and Faraday's constant, respectively; L is the thickness of the membrane or film; and σ_{el} and σ_{ion} are the conductivities of the electrons and oxygen vacancies, respectively.

However, we have previously reported two experimental problems related to Wagner's equation.

Problem 1: The unchanged OCV during electrode degradation [6] [7].

Problem 2: The equilibration process in response to a change in the anode gas

using thick SDC electrolytes [8].

To solve these problems, we proposed the concept of the solvation Hamiltonian of mean force (ϕ), which was discussed by Jarzynski theoretically [3]. After introducing the concept of the solvation Hamiltonian of mean force (ϕ), we noticed that our model was still compatible with Wagner's equation. Thus, we called our model "the modified Wagner's equation".

Using only three graphs, the solvation Hamiltonian of mean force (ϕ) can be explained. The Boltzmann distribution of oxygen ions in the electrolyte at 1073 K is displayed in **Figure 2**. The ions with energies exceeding the ionic activation energy become carriers (hopping ions). **Figure 3** presents an incorrect carrier distribution. According to Jarzynski, an accurate distribution should be a canonical ensemble, as shown in **Figure 4**. The solvation Hamiltonian of mean force (ϕ) can be defined and is illustrated in **Figure 3**.

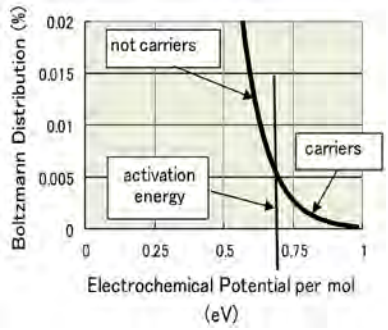


Figure 2. Boltzmann distribution at 1073 K.

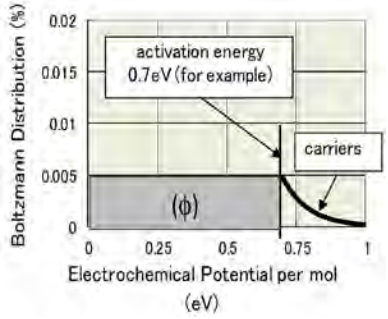


Figure 3. Forbidden distribution of hopping ions.

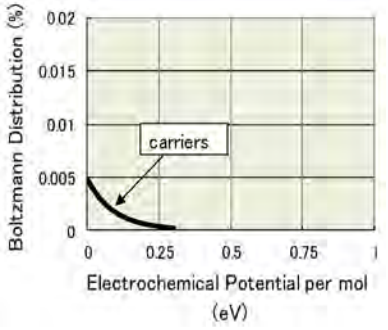


Figure 4. Correct distribution of hopping ions.

Ions with energies exceeding the ionic activation energy are converted into charge carriers (*i.e.*, hopping ions).

According to Jarzynski, an accurate distribution should be a canonical ensemble, and the solvation Hamiltonian of mean force (ϕ) can be defined [3].

The shape of the distribution in this figure should be the same as the shape of the distribution in **Figure 1**.

The electrochemical potential can be separated into the chemical potential and the electrostatic potential, as given by

$$\eta_i = \mu_i + z_i F \varphi \tag{26}$$

where z_i, η_i, μ_i and φ are the valence of species i, the electrochemical potential, the chemical potential and the electrostatic potential, respectively. In Equation (26), the following transformation should be considered during ion hopping:

$$\mu_{i\_\text{hopping}} = \mu_{i\_\text{vacancies}} + NE_a \tag{27}$$

$$z_i F \varphi_{\text{hopping}} = z_i F \varphi_{\text{vacancies}} - NE_a \tag{28}$$

where N, μ_{i_hopping}, $\mu_{i\_\text{vacancies}}$, $\varphi_{i\_\text{hopping}}$, and $\varphi_{i\_\text{vacancies}}$ are Avogadro's number, the chemical potential of hopping ions, the chemical potential of ions in vacancies, the electrical potential of hopping ions, and the electrical potential of ions in vacancies, respectively. For example, E_a is 0.7 eV in SDC electrolytes,

$$\varphi_{\text{hopping}} = \varphi_{\text{vacancies}} + \frac{E_a}{2e} = \varphi_{\text{vacancies}} + 0.35\ \text{V} \tag{29}$$

When there are enough electrons from the anode side in the SDC electrolyte, due to microscopic interactions, this voltage (0.35 V) should be neutralized, and the electrical potential energy of ions should be dissipated by electrons. When the Nernst voltage at 1073 K is 1.15 V, the measured open circuit voltage (OCV) at 1073 K should be 0.80 V (=1.15 V − 0.35 V). Strictly speaking, the voltage loss can be calculated as follows,

$$\text{OCV}_{\text{loss}} = \frac{E_a}{2e} \times \left(1 - t_{\text{ion}}\right) \tag{30}$$

$$t_{\text{ion}} = \frac{\sigma_{\text{ion}}}{\sigma_{\text{el}} + \sigma_{\text{ion}}} \tag{31}$$

where OCV_{loss} and t_{ion} are the voltage loss in the OCV and the ionic transference number, respectively. We conclude that the solvation Hamiltonian of mean force (ϕ) should be

$$\phi = NE_a\,. \tag{32}$$

Therefore, from Equation (27) and Equation (28),

$$\mu_{i\_\text{hopping}} = \mu_{i\_\text{vacancies}} + \phi \tag{33}$$

$$z_i F \varphi_{\text{hopping}} = z_i F \varphi_{\text{vacancies}} - \phi \tag{34}$$

From Equation (30),

$$\Delta E_s = \phi \times \left(1 - t_{\text{ion}}\right) \tag{35}$$

where ΔE_s is the fixed entropy energy difference. Equation (33), Equation (34)

and Equation (35) demonstrate the central concept of "the modified Wagner's equation".

3.3. Consideration of the Compatibility with the Jarzynski's Equality

The compatibility between the Jarzynski's equality and the modified Wagner's equation is discussed in this section. According to Jarzynski [3], the Gibbs volume entropy is given by

$$W \geq \frac{3}{2} NkT \left[\left(\frac{V_A}{V_B} \right)^{\frac{2}{3}} - 1 \right] \geq NkT \ln \left(\frac{V_A}{V_B} \right). \tag{36}$$

In the area of solid-state ionics, the Nernst equation is

$$\mathrm{OCV}_{\mathrm{th}} = \frac{NkT}{4F} \ln \left(\frac{pO_2^A}{pO_2^B} \right), \tag{37}$$

where $\mathrm{OCV}_{\mathrm{th}}$ is the Nernst voltage. Similar to Equation (36) and Equation (37), the study on the OCV in the solid-state ionics is related with the Gibbs volume entropy.

The nonequilibrium work relation is described by Jarzynski's equality [3].

$$\left\langle \mathrm{e}^{-\beta W} \right\rangle = \mathrm{e}^{-\beta \Delta F_T} \tag{38}$$

where β, W and ΔF_T are $1/kT$, the work and the difference in the Helmholtz free energy, respectively. k is the Boltzmann constant. The angular brackets <-> indicate an ensemble over the realizations of the process.

For adiabatic processes, ΔF_T is zero. The following equation should be added [3]:

$$\left\langle \mathrm{e}^{-\beta X} \right\rangle = 1 \tag{39}$$

where X is the energy for the same adiabatic process and ensemble of realizations. In the area of solid-state ionics, X is the ionic activation energy. From Equation (39), an accurate distribution of hopping ions in the electrolytes should be a canonical ensemble. Thus, the Boltzmann distribution during the ion hopping shown in **Figure 4** can be explained by the Jarzynski's equality.

Furthermore, the solvation Hamiltonian of mean force (ϕ) and the fixed entropy energy difference (ΔE_s) in Equation (39) can be defined mathematically [3]. Their mathematical approach is too general. In the area of the generalized second law, the separation of the Boltzmann distribution is called "Maxwell's demon".

Consequently, the concept of "the modified Wagner's equation" can be the specialized version of the Jarzynski's equality.

3.4. Explanation of Our Empirical Equation Using the Modified Wagner's Equation

3.4.1. Assumption from Our Empirical Equation

The explanation of our empirical equation using the modified Wagner's equa-

tion is discussed in this section for when m_{p2} is near m_{p1}. We rewrite Equation (1) and Equation (24) again.

$$\frac{Gm_p}{r_p} \times 1\,\text{kg} = 3kT \tag{40}$$

$$m_{p1}c^2 \times \ln\left(\frac{m_{p2}}{m_{p1}}\right) = 1\,\text{kg} \times G \times \frac{m_{p1}}{r_{p1}} \tag{41}$$

Then, by comparing Equation (40) to Equation (41), we proposed the following equation.

$$m_{p1}c^2 \times \ln\left(\frac{m_{p2}}{m_{p1}}\right) = 3kT \tag{42}$$

When m_{p2} is near m_{p1},

$$\ln\left(\frac{m_{p2}}{m_{p1}}\right) = \frac{m_{p2} - m_{p1}}{m_{p1}} \tag{43}$$

From Equation (42) and Equation (43),

$$\left(m_{p2} - m_{p1}\right)c^2 = 3kT \tag{44}$$

Then, we consider the following a pair production.

$$2h\nu \rightarrow p + \overline{p} \tag{45}$$

where h, ν, p and $\overline{p}$ are the Planck constant, frequency, proton and antiproton, respectively. $h\nu$ does not include both the rest mass and the gravitational potential energy. However, the protons have both the rest mass and the negative gravitational potential energy. Then, we proposed the following equation.

$$m_p c^2 = h\nu - 3kT \tag{46}$$

$$\phi = 3kT\,. \tag{47}$$

Equation (46) and Equation (47) are hinted from Equation (32) and Equation (34). The electrical potential energy (internal energy) of hopping ions should be smaller than that of ions in vacancies. The value of the decrease is equal to the solvation Hamiltonian of mean force (ϕ).

This means that the rest mass should decrease from $h\nu$. $3kT$ (T = 3.20 K) is 0.0083 eV, which is much smaller than the energy of the rest mass of protons (938 MeV). The increasing ratio is 8.8×10^{-13}, and the decrease in the rest mass cannot be observed easily.

3.4.2. Consideration for the Temperature of the Proton

The solvation Hamiltonian of mean force (ϕ) is rewritten.

$$\phi = 3kT\,. \tag{48}$$

Here, T is 3.2 K, which is similar to the temperature of the cosmic microwave background of 2.73 K. This does not mean that the temperature of the proton should be near 9 K.

For example, in the area of solid-state ionics, the solvation Hamiltonian of mean force (ϕ) should be equal to the ionic activation energy. When the ionic activation energy is 1.0 eV, the calculated temperature is 1.16×10^4 K. This temperature is different from the background temperature of 1073 K. $3kT$ is the interaction energy between the proton and the rest mass.

The proton consists of three quarks. Thus, the solvation Hamiltonian of mean force (ϕ) of one quark is

$$\phi_{\text{quark}} = kT \,. \tag{49}$$

Equation (49) is more reasonable than Equation (48). The temperature of the cosmic microwave background of 2.73 K is the universal temperature. The origin of this temperature may be from the interaction energy between the quark and gravity.

3.4.3. The Lower Limit of the Rest Mass

The particle, which has a total energy less than 0.0083 eV ($3kT$), cannot be influenced by gravity. For example, in the area of solid-state ionics, ions that have an energy less than the ionic activation energy cannot be carriers. The lower limit of the small mass is 0.0083 eV ($3kT$). The mass of a neutrino is assumed to be less than 0.5 eV. This may provide a solution for the neutrino oscillation problem.

From Equation (35), inevitable dissipation was predicted. When the transference of gravity is not exactly 1, there should be inevitable dissipation. The interaction coefficient should be on the order of 8.8×10^{-13}. This value is similar to the value of the weak interaction.

3.4.4. Consideration for Einstein's Equivalence Principle

Equation (46) is rewritten.

$$m_p c^2 = h\nu - 3kT \tag{50}$$

In Equation (50), the value of m_{p1} can be changed. Then, by measuring the decreasing ratio of the rest mass of the proton after a pair is produced in the falling elevator, we can measure the gravitation potential in the elevator. However, according to Einstein, not everybody cannot do it. To solve this problem, we proposed the following equation.

$$\frac{m_{p1}}{m_{p2}} = \frac{k_1}{k_2} \tag{51}$$

where k_1 and k_2 are the Boltzmann constant in place 1 and in place 2. The Boltzmann constant was defined from the mass of a proton. Thus, Equation (51) is correct. Equation (50) should be

$$h\nu = m_{p1} c^2 + 3k_1 T \tag{52}$$

Then, by using $k_1 T$ instead of kT, in the following Equation (53), the calculated value was constant. Thus, we cannot measure the gravitation potential in the elevator. This means that $3kT$ should increase with increasing m_p. Conse-

quently, the solvation Hamiltonian mean force (ϕ) should increase with increasing distance.

$$\frac{3k_1T}{m_{p1}c^2} = 8.8\times10^{-13} \tag{53}$$

3.4.5. Consideration When m_{p2} Is Not near m_{p1}

When m_{p2} is not near m_{p1}, the explanation for our empirical equation using the modified Wagner's equation is discussed in this section. Equation (42) is rewritten.

$$m_{p1}c^2 \times \ln\left(\frac{m_{p2}}{m_{p1}}\right) = 3kT \tag{54}$$

Thus,

$$\ln\left(\frac{m_{p2}}{m_{p1}}\right) = \frac{m_{p1}c^2}{3kT} \tag{55}$$

Thus,

$$\frac{m_{p2}}{m_{p1}} = \mathrm{e}^{\frac{m_{p1}c^2}{3kT}} \tag{56}$$

Then, k is a function of the mass of a proton (m_p) and cannot be constant. In Equation (56), the relationship between the rest mass and the distance cannot be obtained. It is impossible to progress any further. This conclusion is the same as the conclusion proposed by Yasuo Katayama in Equation (16).

Using Ted Jacobson's theory, this problem can be solved.

4. Explanation for Our Model for Gravity

4.1. Ted Jacobson's Theory

Ted Jacobson discovered that the general relative theory is governed by the following Equation (57).

$$\mathrm{d}S = \frac{\partial S}{\partial E}\mathrm{d}E + \frac{\partial S}{\partial V}\mathrm{d}V \tag{57}$$

where S, E and V are the entropy, the internal energy and the volume, respectively.

4.2. Our Model for Gravity

Equation (46) is rewritten.

$$m_p c^2 = h\nu - 3kT \tag{58}$$

In Equation (58), internal energy (E) should decrease. However, the entropy should not decrease. From Equation (57), the Gibbs volume entropy (S_{Gibbs}) can be defined.

$$S_{\mathrm{Gibbs}}\left(m_p c^2\right)\times T = S_{\mathrm{Gibbs}}\left(h\nu\right)\times T + 3kT \tag{59}$$

Then, we must consider the degree of freedom (x, y and z direction). In Equation (59), the Gibbs volume entropy (S_{Gibbs}) for one direction is

$$S_{\text{Gibbs}}\left(m_p c^2\right) = S_{\text{Gibbs}}\left(h\nu\right) + kT \tag{60}$$

Next, dV in the distance (r) from the mass should be

$$\mathrm{d}V = 4\pi r^2 \times \mathrm{d}r \tag{60}$$

Thus,

$$\frac{\partial S_{\text{Gibbs}}}{\partial V}\mathrm{d}V = \frac{S_{\text{Gibbs}}}{\frac{4}{3}\pi r^3} \times 4\pi r^2 \times \mathrm{d}r = \frac{3S_{\text{Gibbs}}}{r}\mathrm{d}r \tag{61}$$

Therefore, from Equation (60) and Equation (61),

$$\frac{\mathrm{d}S_{\text{Gibbs}}}{\mathrm{d}r}T = \frac{3kT}{r} \tag{62}$$

When the number of protons is n,

$$\frac{\mathrm{d}S_{\text{Gibbs}}}{\mathrm{d}r}T = \frac{3nkT}{r} \tag{63}$$

It is important that the direction is opposite to that of the gravitational potential. However, Equation (63) is similar to Newton's Law. k is one of the "defining constants" and becomes part of the 2019 redefinition of SI base units. However, from Equation (53), the value of k should be changed. Equation (53) is rewritten.

$$\frac{3k_1 T}{m_{p1}c^2} = 8.8\times 10^{-13} \tag{64}$$

It is possible that the increase in the rest mass can be observed on the galaxy scale. Then, k should increase with increasing the mass of a proton. This may provide a solution for the dark matter problem.

4.3. Consideration for the Origin of the Solvation Hamiltonian of Mean Force (ϕ) in the Universe

The solvation Hamiltonian mean force (ϕ) is assumed to originate from the asymmetry of space-time. The asymmetry of space-time was already discovered and is known as "chirality". According to Einstein,

$$x^2 + y^2 + z^2 - ct^2 = 0 \tag{65}$$

We proposed the following equation.

$$Z_1^2 + Z_2^2 + Z_3^2 + Z_0^2 = 0 \tag{65}$$

where Z_1, Z_2, Z_3 and Z_0 are complex numbers. Then, the asymmetry $e^{i\theta}$ can be defined. In the area of solid-state ionics, the asymmetry can be produced artificially. For example, YSZ (yttria stabilized zirconia) can be obtained from ZrO_2 with an addition of 8% Y_2O_3. Then, ions can move in the YSZ electrolytes due to the existence of vacancies. Therefore, the ionic activation energy (E_a) can be defined for the ionic conduction. We noticed that the ionic activation energy can be defined as the solvation Hamiltonian mean force (ϕ). Then, the Gibbs volume

entropy can be measured experimentally. Consequently, we proposed the fifth dimension, which is the internal dimension originating from the asymmetry that acts to separate the Boltzmann distribution.

5. Conclusion

We discovered an empirical relationship. Then, we examined Yasuo Katayama's theory. Previously, we developed "the modified Wagner's concept" from our own experimental results. We attempted to explain this concept from the generalized second law. The concept of "the modified Wagner's equation" can be the specialized version of the Jarzynski's equality. We examined Ted Jacobson's theory, which is very difficult. However, the basic concept is very easy for us. The Gibbs volume entropy can be defined. Using the Gibbs volume entropy, we tried to explain our empirical equation and tried to comprehensively develop Yasuo Katayama's theory. The calculated temperature was 3.20 K, which is similar to the temperature of the cosmic microwave background of 2.73 K and was much larger than the Unruh temperature. Several topics were suggested for future discussion.

Conflicts of Interest

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

References

[1] Jacobson, T. (1995) *Physical Review Letters*, **75**, 1260-1263. arXiv:gr-qc/9504004 https://doi.org/10.1103/PhysRevLett.75.1260

[2] Verlinde, E.P. (2011) *Journal of High Energy Physics*, **2011**, Article Number: 29. arXiv:1001.0785 https://doi.org/10.1007/JHEP04(2011)029

[3] Jarzynski, C. (2017) *Physical Review X*, **7**, 011008. https://doi.org/10.1103/PhysRevX.7.011008

[4] Miyashita, T. (2020) Open-Circuit Voltage Anomalies in Yttria-Stabilized Zirconia and Samaria-Doped Ceria Bilayered Electrolytes. https://ecsarxiv.org/xhn73/ https://doi.org/10.1149/osf.io/xhn73

[5] Katayama, Y. http://home.catv.ne.jp/dd/pub/pot/pot.html#0

[6] Miyashita, T. (2006) *Journal of Materials Science*, **41**, 3183. https://doi.org/10.1007/s10853-006-6371-8

[7] Miyashita, T. (2011) *Electrochemical and Solid State Letters*, **14**, 1. https://doi.org/10.1149/1.3581208

[8] Miyashita, T. (2017) *Journal of the Electrochemical Society*, **164**, 3190. https://doi.org/10.1149/2.0251711jes

Journal of Modern Physics, 2020, 11, 1193-1214
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Group Geometric Algebra and the Standard Model

Carl Brannen

Retired, Pullman, WA, USA
Email: carl.brannen@wsu.edu

How to cite this paper: Brannen, C. (2020) Group Geometric Algebra and the Standard Model. *Journal of Modern Physics*, **11**, 1193-1214.
https://doi.org/10.4236/jmp.2020.118075

Received: July 3, 2020
Accepted: August 14, 2020
Published: August 17, 2020

Abstract

We show how to generalize the Weyl equation to include the Standard Model fermions and a dark matter fermion. The 2 × 2 complex matrices are a matrix ring R. A finite group G can be used to define a group algebra $G[R]$ which is a generalization of the ring. For a group of size N, this defines N Weyl equations coupled by the group operation. We use the group character table to uncouple the equations by diagonalizing the group algebra. Using the full octahedral point symmetry group for G, our uncoupled Weyl equations have the symmetry of the Standard Model fermions plus a dark matter particle. We describe the symmetry properties of dark matter.

Keywords

Weyl Equation, Dark Matter, Standard Model

1. Introduction

We will write the Weyl equation [1] for a single massless fermion as

$$\sigma^{\mu}\partial_{\mu}\psi = I_2\frac{\partial\psi}{\partial t} \pm \sigma_x\frac{\partial\psi}{\partial x} \pm \sigma_y\frac{\partial\psi}{\partial y} \pm \sigma_z\frac{\partial\psi}{\partial z} = 0, \quad (1)$$

where I_2 is the 2 × 2 unit matrix, the σ_j is the Pauli spin matrices. The plus signs are for the left handed Standard Model fermions and the minus signs are for the right handed fermions. In the above, the Pauli algebra has a basis of four 2 × 2 complex matrices $\{\sigma^{\mu}\} = \{\sigma_t, \sigma_x, \sigma_y, \sigma_z\} = \{I_2, \sigma_x, \sigma_y, \sigma_z\}$ where each of the four σ^{μ} has two nonzero complex elements. The wave function ψ is a 1 × 2 vector so the Weyl equation consists of two coupled differential equations with four terms in each. We will generalize the σ^{μ} to $\sigma^{\mu g}$ so that they depend on a group element g. There will now be $4N$ different $\sigma^{\mu g}$ where N is the number of elements in the finite group G. These matrices are still linearly

independent in the sense that

$$\Sigma_{\mu,g} a_{\mu g} \sigma^{\mu g} = 0 \tag{2}$$

implies that the $4N$ complex constants $a_{\mu g}$ are all zero.

Changing the signs of the Pauli spin matrices, as seen above in the left and right handed Weyl equations, can be thought of as a symmetry operation. The symmetry group has two elements, it is the point group symmetry that Hestenes and Holt [2] call $\overline{22}$. We will label the point group symmetries we use with their notation and to avoid confusion with numerals will prepend them with "Geo" as in $\text{Geo}\overline{22}$. The groups explored in this paper are the full octahedral group Geo43, some of its subgroups and a tripled Geo43 with $48\times 3 = 144$ elements we will call 3Geo43. An element of Geo43 or its subgroups can be described by its effect on the Pauli spin matrices. In general, the three matrices can be permuted and can be independently negated. Together, these give the $3!\times 2^3 = 48$ elements of Geo43.

The finite point group $\text{Geo}\overline{22}$ is a good example to make clear how we are generalizing the Weyl equation. First, to make our notation concise and descriptive, let's define the two elements of the group as $\text{Geo}\overline{22} = \{(+),(-)\}$ so that the multiplication rules are $(+)(+)=(+)$, $(+)(-)=(-)$, $(-)(+)=(-)$ and $(-)(-)=(+)$. Then the eight generalizations of the four Pauli spin matrices are $\sigma^{\mu\pm}$. And the wave function ψ is generalized to a wave function with twice as many degrees of freedom $\psi^{\pm}$.

An obvious way to use the $\sigma^{\mu\pm}$ to generalize the Weyl equation is to write

$$\sigma^{\mu\pm}\partial_{\mu}\psi^{\pm} = 0. \quad \text{[Not what we will use!]} \tag{3}$$

This would give two independent Weyl equations, one for each sign. In total there will be four differential equations with four terms in each. But in this generalization the group G has no effect on the equations other than to increase the particle content by a factor N. Instead, we will couple the group designation g for the Pauli spin matrix $\sigma^{\mu g}$ with the designation, say h for the wave function ψ^{h}. Then the positive and negative Weyl equations are:

$$\begin{aligned} &\sigma^{\mu+}\partial_{\mu}\psi^{+} + \sigma^{\mu-}\partial_{\mu}\psi^{-} = 0, \quad [+\text{equation}] \\ &\sigma^{\mu-}\partial_{\mu}\psi^{+} + \sigma^{\mu+}\partial_{\mu}\psi^{-} = 0. \quad [-\text{equation}] \end{aligned} \tag{4}$$

The top, +, equation corresponds to the two ways of obtaining (+) in the group, that is, $(+)=(+)(+)=(-)(-)$, while the bottom equation gives the two ways of obtaining $(-)=(-)(+)=(+)(-)$. There are now two cross coupled Weyl equations. In total these are four differential equations with eight terms each. Our task will be to manipulate these equations so as to decouple them into uncoupled Weyl equations.

To model the Standard Model fermions, we will use a group of size $N=144$. This will give 2 × 144 differential equations with 4 × 144 terms in each for a total of $8\times 144^2 = 165888$ terms. It would be difficult to decouple such a complicated problem by hand. We will use methods developed in the 19th century known

now as "harmonic analysis". [3] This was a method of solving wave equation problems based on symmetry. Originally this required a model of the medium on which the waves acted but this was abstracted away to give the irreducible representations of symmetry now popular in elementary particle theory. A readable introduction to the development is given in [4]. For linear symmetry, harmonic analysis prescribed using the Fourier transform to convert the waves into sines and cosines. For situations with spherical symmetry the same type of transform converts to spherical harmonics which can be abstracted to give the irreducible representations of SO(3). Harmonic analysis also applies to situations with discrete symmetry. For Abelian finite symmetry groups the Fourier transform becomes the "discrete Fourier Transform". In this paper we will expand the discrete Fourier transform to situations with non Abelian symmetry groups [5] and use this to solve the decoupling problem.

To write the equivalent of Equation (4) for more general groups, we need a way to describe all the possible group products that give a product g. We then sum over all these, assigning the first group product to the σ and the second to the ψ. Since $(g/h)h = (gh^{-1})h = g$, we can sum over h and use:

$$\Sigma_h \sigma^{\mu g/h} \partial_\mu \psi^h = 0 \tag{5}$$

for each of the $g \in G$. This gives the g part of the wave equation and considering the N elements of the group we have N coupled Weyl equations. The primary work of this paper is to uncouple equations like these.

In the Weyl representation [1] the 4 × 4 gamma matrices are:

$$\gamma^0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}. \tag{6}$$

Dirac's wave equation is:

$$i\gamma^\mu \partial_\mu \psi = m\psi, \tag{7}$$

but if we multiply both sides by γ^0 and rewrite it using the 2 × 2 Pauli matrices, and split the 4-vector ψ into two 2-vectors ψ_L and ψ_R we have

$$\left(\partial_t + \begin{pmatrix} -\sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \partial_k\right) \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = -im \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix}. \tag{8}$$

The above are two Weyl equations with ψ_L using negative Pauli spin matrices and ψ_R using the positive, but with the equations coupled on the right by the mass term. Using the $\sigma^{\mu g}$ notation these two coupled Weyl equations are:

$$\begin{aligned} \sigma^{\mu -} \partial_\mu \psi_L &= -im\psi_R, \\ \sigma^{\mu +} \partial_\mu \psi_R &= -im\psi_L. \end{aligned} \tag{9}$$

Considering the above as arising from the Dirac equation, the $\sigma^{\mu \pm}$ are 4 × 4 matrices with nonzero entries only in the top left 2 × 2 block (for the case) or the bottom right 2 × 2 block (for the + case). Setting the mass to zero gives two uncoupled Weyl equations. Physically, these two Weyl equations define two

massless particles that travel at the speed of light, a limit that cannot be achieved in the physical world. An alternative method of decoupling ψ_L and ψ_R is to require that one of them be zero. Then the equation satisfied by the other is the massless Weyl equation.

In quantum information theory, a spin-1/2 state is represented by a complex 2-vector so there is no momentum and the energy is directly proportional to the mass and is therefore an uninteresting constant. We can get the quantum information version of the Dirac equation by defining ψ_+ and ψ_- by $\psi_\pm = (\psi_R \pm \psi_L)/2$. As in the alternative derivation of the massless Weyl equations, we set one of the $\psi_\pm$ to be zero and see what equations the other satisfies. Putting $\psi_- = 0$ and multiplying the equations by 2, we obtain the equations for ψ_+ when $\psi_- = 0$:

$$\begin{aligned} \sigma^{\mu-}\partial_\mu \psi_+ &= -im\psi_+, \\ \sigma^{\mu+}\partial_\mu \psi_+ &= -im\psi_+. \end{aligned} \tag{10}$$

Expanding $\sigma^{\mu\pm}\partial_\mu = \partial_t \pm \sigma_j \partial_j$ into time and spatial parts, taking the sum and difference between the two equations, and dividing by 2 we get:

$$\begin{aligned} \partial_t \psi_+ &= -im\psi_+, \\ \sigma_j \partial_j \psi_+ &= 0. \end{aligned} \tag{11}$$

The first equation defines the energy as constant while the second defines the momentum as zero. These are the wave equations for a stationary spin-1/2 particle of mass *m*. We get the same result for ψ_- but since this is a difference between ψ_L and ψ_R we can identify this as the antiparticle. Thus the quantum information limit of the Dirac equation includes a particle and an antiparticle.

In this paper we will be further extending this quantum information limit to include all the fermions of the Standard Model [6] [7]. We will be using irreducible representations of symmetry and these will naturally mix ψ_L and ψ_R as either $\psi_R + \psi_L$ or $\psi_R - \psi_L$ and we will treat these respectively as the particle and antiparticle. [8]

Unlike ψ_L and ψ_R, the $\psi_\pm$ states have definite electric charge (that is, in a physical measurement of electric charge they are eigenstates of the physical electric charge operator). This is in contrast to the quantum field theory assumption of the Standard Model where antiparticles are treated as having the same charge as particles, but travelling backwards in time so the measured charge is negated. Another way of describing this difference is that in this paper, we will be treating all Standard Model fermions as travelling forwards in time, and we will be treating mass as a slightly messy interaction between particle and antiparticle rather than a simpler interaction between left and right handed states.

We will illustrate the uncoupling of the generalized Weyl equation in three steps. The first step will use the "Geo$\overline{22}$" point group with two elements, the identity (+) and the inversion (−), and is the subject of Section 1.0. We write the

generalized Weyl equations which are four coupled partial differential equations, and decouple them into two Weyl equations corresponding to two particles. The decoupling is equivalent to putting the group algebra $\overline{22}\left[\sigma^{\mu}\right]$ into block diagonal form. A short cut to diagonalizing a group algebra is to use the finite group's character table. In general, an Abelian group of size N has a character table of size $N \times N$, and an Abelian group generated by a single element has a character table whose entries are the same as the discrete Fourier transform. Later we generalize the character tables of non Abelian symmetries to also be of size $N \times N$ so they can be used as a non Abelian generalization of the discrete Fourier transform.

The second step will use the point group Geo3 that corresponds to the six permutations on the three Pauli spin matrices and is discussed in Section 2.0. The group's generalized Weyl equation has twelve coupled differential equations with 24 terms each. We will use the character table for Geo3 to assist in decoupling the equations. The group is small enough that writing down the decoupled Weyl equations is manageable. Since Geo3 is not Abelian, block diagonalizing $3\left[\sigma^{\mu}\right]$ leaves a 2×2 block on the diagonal. While the group size of Geo3 is 6, the character table has only 3 columns and 3 rows but we show how to expand it to a 6×6 table that can be interpreted as a generalization of the discrete Fourier transform to the non Abelian Geo3 symmetry group. The expanded character table defines an internal symmetry that is an SU(2) analog of the color SU(3) internal symmetry of quarks. Note that the matrices of the Weyl and Dirac wave equations transform the way that density matrices transform rather than state vectors. This is significant because our generalized discrete Fourier transform is in the manner of a state vector transform, but it is on an object that transforms as a density matrix. We discuss the thermodynamics of internal symmetries.

In Section 3.0 we discuss the full octahedral group Geo43. We use the character table methods shown in the previous section to read off the result of uncoupling. We find four Standard Model first generation leptons, *i.e.* electron, positron, neutrino and anti-neutrino, four Standard model first generation quarks and a set of Weyl equations with internal SU(2) symmetry left over for dark matter and anti-dark matter. Thinking of our transformation as a discrete Fourier transform, we can reverse the process and act on the electric charge, weak hypercharge and weak isospin to find what elements of $43\left[\sigma^{\mu}\right]$ transform to these operators. The result is compatible with the observation that Fourier transforms are useful for simplifying complicated structures in physics.

Section 4.0 briefly discusses how we can triple the Geo43 point group to give the three generations of the Standard Model.

In Section 5.0 we discuss the dark matter doublet found in the previous sections and the implications for dark matter. The internal SU(2) symmetry is assumed to act similarly to the internal SU(3) color symmetry so we call them "dark quarks" or "duarks". Corresponding to the quark color force boson the gluon, dark quarks are bound by a dark color force boson called the "duon". We

assume duarks combine similarly to the usual quarks so we have dark matter composed of "daryons" and "desons". Instead of colors that sum to white, the dark quarks do not interact with photons so their colors must sum to black. Since they are SU(2) instead of SU(3), we need only two dark colors to form a basis set for their state vectors and we call the two dark colors "doom" and "gloom".

We finish the paper with a brief acknowledgement.

2. $\overline{22}\left[\sigma^{\mu}\right]$, Generalized Weyl Equations and Decoupling

The group $\mathrm{Geo}\overline{22}$ is equivalent to S_2, the group of permutations of two elements. A common notation for the two elements of the group is the permutation notation but instead we will use ±:

$$\begin{aligned} (\,) &= (+), \\ (12) &= (-), \end{aligned} \tag{12}$$

and leave off the parentheses when the meaning remains clear. The (−) elements changes the signs of the Pauli spin matrices and the identity (+) leaves the signs unchanged. The multiplication table is:

$$\begin{array}{c|cc} \overline{22} & (+) & (-) \\ \hline (+) & (+) & (-) \\ (-) & (-) & (+) \end{array} \tag{13}$$

The four basis elements of the Pauli algebra are $\left\{I_2, \sigma_x, \sigma_y, \sigma_z\right\}$; The $\mathrm{Geo}\overline{22}$ group has two elements so the group algebra $\overline{22}\left[\sigma^{\mu}\right]$ will have two basis elements for each of these, corresponding to the two group elements (+) and (−): $\left\{I_2^+, \sigma_x^+, \sigma_y^+, \sigma_z^+ I_2^-, \sigma_x^-, \sigma_y^-, \sigma_z^-\right\}$. These eight elements are a basis for the $\overline{22}\left[\sigma^{\mu}\right]$ algebra so that an arbitrary element α of the algebra can be written as

$$\alpha = \alpha_{I+} I_2^+ + \alpha_{x+}\sigma_x^+ + \alpha_{y+}\sigma_y^+ + \alpha_{z+}\sigma_z^+ + \alpha_{I-} I_2^- + \alpha_{x-}\sigma_x^- + \alpha_{y-}\sigma_y^- + \alpha_{z-}\sigma_z^- \tag{14}$$

where the $\alpha_{\chi\pm}$ are eight complex numbers.

Addition is the usual for a complex vector space with 8 components, as is multiplication by a complex number. We will discuss statistical mechanics of group algebra quantum states in the next section; that will be our only use of multiplication of two group algebra elements so we define it here. Multiplication is term by term with products of basis elements given by the $\mathrm{Geo}\overline{22}$ group product and the usual Pauli algebra rules. The multiplicative identity for the group algebra is I_2^+. So for example:

$$\begin{aligned} &(1+2i)\sigma_x^+ (2-3i)\sigma_y^- = (2+4i-3i+6) i\sigma_z^{+-} = (-1+8i)\sigma_z^-, \\ &I_2^+ \alpha = \alpha I_2^+ = \alpha, \\ &\left(\sigma_x^-\right)\left(\sigma_x^- + \sigma_y^+\right) = I_2^+ + i\sigma_z^-, \end{aligned} \tag{15}$$

where we've used the usual Pauli algebra multiplication $\sigma_x\sigma_y = i\sigma_z$, $\sigma_x\sigma_x = I_2$ and the $\mathrm{Geo}\overline{22}$ group multiplication rules $(+)(+) = (-)(-) = (+)$,

$(+)(-)=(-)(+)=(-)$. These rules reduce any product to the $\overline{22}\left[\sigma^{\mu}\right]$ basis.

The Weyl equation for $\overline{22}\left[\sigma^{\mu}\right]$ is

$$\sigma^{\mu g/h}\partial_{\mu}\Psi^{h}=0, \tag{16}$$

where h is summed over $\mathrm{Geo}\overline{22}=\{(+),(-)\}$. There are two equations, one for $g=(+)$, the other for $g=(-)$. Using $1/h=h$ for $\mathrm{Geo}\overline{22}$ so that $g/h=gh$ and writing out the sums over h gives two equations, the first for $g=(+)$ and the second for $g=(-)$ as follows:

$$\begin{aligned}\sigma^{\mu(+)(+)}\partial_{\mu}\Psi^{+}+\sigma^{\mu(+)(-)}\partial_{\mu}\Psi^{-}&=0,\\ \sigma^{\mu(-)(-)}\partial_{\mu}\Psi^{-}+\sigma^{\mu(-)(+)}\partial_{\mu}\Psi^{+}&=0.\end{aligned} \tag{17}$$

Using the $\mathrm{Geo}\overline{22}$ group multiplication, we reduce the *gh* products such as (+)(+) to give

$$\begin{aligned}\sigma^{\mu+}\partial_{\mu}\Psi^{+}+\sigma^{\mu-}\partial_{\mu}\Psi^{-}&=0,\\ \sigma^{\mu+}\partial_{\mu}\Psi^{-}+\sigma^{\mu-}\partial_{\mu}\Psi^{+}&=0,\end{aligned} \tag{18}$$

and we see we have two coupled Weyl equations.

To uncouple these two equations, consider the character table for the $\mathrm{Geo}\overline{22}$ group:

Class:	$\{(+)\}$	$\{(-)\}$
Size:	1	1
A	1	+1
B	1	−1

(19)

The group is Abelian so the two classes have only a single element in each. There are two irreducible representations consisting of the sum and difference of the (+) part and (−) part. Taking this as a hint for how to decouple the two Weyl equations, we compute the sum and differences of Equation (18) and indeed obtain two decoupled Weyl equations:

$$\begin{aligned}\left(\sigma^{\mu+}+\sigma^{\mu-}\right)/2\,\partial_{\mu}\left(\psi^{+}+\psi^{-}\right)&=0,\\ \left(\sigma^{\mu+}-\sigma^{\mu-}\right)/2\,\partial_{\mu}\left(\psi^{+}-\psi\right)^{-}&=0.\end{aligned} \tag{20}$$

The normalization factor of 1/2 is included so that the decoupled Pauli basis elements $\tau^{\mu\pm}=\left(\sigma^{\mu+}\pm\sigma^{\mu-}\right)/2$ satisfy the equation $\left(\tau^{\mu\pm}\right)^{2}=I_{2}^{+}$, which corresponds to the usual Pauli algebra equation $\left(\sigma_{\mu}\right)^{2}=1$.

We can also illustrate the method using a matrix representation of the finite group $\mathrm{Geo}\overline{22}$. Using

$$(+)=\begin{pmatrix}1&0\\0&1\end{pmatrix},\ (-)=\begin{pmatrix}0&1\\1&0\end{pmatrix}, \tag{21}$$

as a representation, note that the representation satisfies the group multiplication and that they are linearly independent matrices. That is, if $\alpha_{+}(+)+\alpha_{-}(-)=0$ for complex numbers $\alpha_{\pm}$, then $\alpha_{\pm}=0$. The four basis elements of the Pauli algebra σ^{μ} are four 2 × 2 complex matrices that are

linearly independent. Similarly, the eight basis elements of the $\mathrm{Geo}\overline{22}\left[\sigma^{\mu}\right]$ are also linearly independent.

In an abuse of notation, we can write the four Pauli algebra matrices as a single 2 × 2 complex matrix:

$$\sigma^{\mu}=\begin{pmatrix} I_2+\sigma_z & \sigma_x-i\sigma_y \\ \sigma_x+i\sigma_y & I_2-\sigma_z \end{pmatrix}, \tag{22}$$

that defines the four 2 × 2 complex matrices. The same abuse of notation allows us to write the eight basis elements of $\overline{22}\left[\sigma^{\mu}\right]$ in a single 4 × 4 matrix:

$$\sigma^{\mu\pm}=\begin{pmatrix} I_2^{+}+\sigma_z^{+} & \sigma_x^{+}-i\sigma_y^{+} & I_2^{-}+\sigma_z^{-} & \sigma_x^{-}-i\sigma_y^{-} \\ \sigma_x^{+}+i\sigma_y^{+} & I_2^{+}-\sigma_z^{+} & \sigma_x^{-}+i\sigma_y^{-} & I_2^{-}-\sigma_z^{-} \\ I_2^{-}+\sigma_z^{-} & \sigma_x^{-}-i\sigma_y^{-} & I_2^{+}+\sigma_z^{+} & \sigma_x^{+}-i\sigma_y^{+} \\ \sigma_x^{-}+i\sigma_y^{-} & I_2^{-}-\sigma_z^{-} & \sigma_x^{+}+i\sigma_y^{+} & I_2^{+}-\sigma_z^{+} \end{pmatrix}. \tag{23}$$

Note that in the above, the (+) and (−) supercripts follow the $\mathrm{Geo}\overline{22}$ group representation of Equation (21) while the Pauli algebra basis of Equation (22) appears in the subscripts of each of the four 2 × 2 blocks. As an example, the above indicates that the identity I_2^{+} is the 4×4 identity matrix, and that

$$\sigma_x^{-}=\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{24}$$

etc. Note that our 8 basis elements only describe half the degrees of freedom in a 4 × 4 complex matrix. We will now transform this representation to 2 × 2 block diagonal.

Given a matrix U and its inverse U^{-1} we can transform an algebra by $\alpha\to\alpha'=U\alpha U^{-1}$. Choosing

$$U=U^{-1}=\frac{1}{\sqrt{2}}\begin{pmatrix} I_2 & I_2 \\ I_2 & -I_2 \end{pmatrix} \tag{25}$$

conveniently puts the $\overline{22}\left[\sigma^{\mu}\right]$ group algebra into block diagonal form. This corresponds to a transformation on our representation of $\mathrm{Geo}\overline{22}$. Using the usual abuse of notation, in an even more confusing way, that is, to give one equation that shows the transformation of both the (+) and (−) elements of $\mathrm{Geo}\overline{22}$, we have:

$$U\begin{pmatrix} (+) & (-) \\ (-) & (+) \end{pmatrix}U^{-1}=\frac{1}{2}\begin{pmatrix} (+)+(-) & 0 \\ 0 & (+)-(-) \end{pmatrix}. \tag{26}$$

This transformation puts the $\overline{22}\left[\sigma^{\mu}\right]$ basis elements into block diagonal form. This is similar to the Weyl choice of representation for the gamma matrices which puts the $\gamma^0\gamma^{\mu}$ matrices into block diagonal form. Then putting the mass zero uncouples the Dirac equation into two Weyl equations.

In the next two sections we will be decoupling Weyl equations for more complicated finite groups; we will approach those problems as a matter of block

diagonalization of the finite group just as here we have block diagonalized $\text{Geo}\overline{22}$.

The two coupled Weyl equations of Equation (20) differ in the wave function: $\psi^+ \pm \psi^-$. This fact can be extracted from the block diagonal form of $\overline{22}\left[\sigma^\mu\right]$ given in Equation (26) by considering it as the symmetry of a density matrix. The $U\alpha U^{-1}$ transformation of the algebra is also how density matrices are traditionally transformed. This follows from the fact that the Weyl equations are matrix equations and so transform as matrices.

3. $3\left[\sigma^\mu\right]$, Nonabelian Group Algebras and Internal Symmetry

The Geo3 point group has six elements. The group is equivalent to the permutation group on three elements, S_3. Previously we had considered $\text{Geo}\overline{22}$ as a transformation on the Pauli spin matrices that negated all of their signs. Similarly, we can consider Geo3 as a transformation on the Pauli spin matrices that is a permutation on the three spatial components σ_x, σ_y and σ_z. Representing the spin matrices by X, Y, and Z and defining their permutations by how these are ordered, the conversion from the usual permutation notation is:

$$\begin{aligned} &( ) \equiv XYZ, \quad (123) \equiv YZX, \quad (132) \equiv ZXY, \\ &(12) \equiv YXZ, \quad (13) \equiv ZYX, \quad (23) \equiv XZY. \end{aligned} \tag{27}$$

Later we will consider the Geo43 point group which can be considered as all the six permutations of the Pauli spin matrices along with any individual negations where there are $2^3 = 8$ possibilities. The Geo43 group therefore has $6 \times 2^3 = 48$ elements.

Since the Pauli algebra has a basis of four elements, the basis for the $3\left[\sigma^\mu\right]$ algebra has $4 \times 6 = 24$ elements and we will designate them as $\left\{I_2^{(\,)}, I_2^{(123)}, \cdots, \sigma_z^{(23)}\right\}$.

In the previous section we were able to completely diagonalize the $\text{Geo}\overline{22}$ part of the $\overline{22}\left[\sigma^\mu\right]$ algebra. This was possible only because $\text{Geo}\overline{22}$ is Abelian. Since Geo3 is non Abelian we will be unable to completely diagonalize the Geo3 part of the $3\left[\sigma^\mu\right]$ algebra.

A group algebra is a generalization of a field or ring. Here we've been using the Pauli algebra or 2 × 2 complex matrices as the ring and we've been designating this ring as "σ^μ". The mathematics literature on group algebras frequently does not specify the field or ring and instead consider the general subject of a group algebra over an unspecified field that satisfies a few easy requirements (that are met by the Pauli algebra). Given the arbitrariness of the choice of ring or field, a more convenient choice is the field of complex numbers. The group algebra over the field of complex numbers is discussed at length in Hammermesh's classic book "Group Theory and its Applications to Physical Problems" Section 3-17 [9] where he uses the group algebra as a way of defining

all the irreducible representations of a finite symmetry. Here we will be doing the same diagonalization as Hammermesh but with the purpose of uncoupling the generalized Weyl equations.

Each line of the character table of the finite group corresponds to a diagonal block. The size of the block is defined by the character of the identity, here called ( ). Our Geo$\overline{22}$ example of the previous section was Abelian so all the representations have character 1 for the identity and the block diagonalized group algebra had only 1×1 blocks.

The character table for Geo3 has three irreducible representations, two, A and B, with character 1 for the identity, and the third C with character 2 for the identity as shown in the column labeled "{( )}":

Class :	$\{(\ )\}$	$\{(123)\}$	$\{(12)\}$
Size :	1	2	3
A	1	+1	+1
B	1	+1	−1
C	2	−1	0

(28)

This corresponds to two 1×1 diagonal blocks for the A and B irreps, and one 2×2 diagonal block for the C irrep.

As elements of the algebra, irreducible representations commute with everything in the algebra. As Hammermesh discusses [9], the A and B irreps are zero outside of their own 1×1 blocks, and similarly the C irrep is a multiple of the 2×2 identity for its block. With our usual abuse of notation:

$$\sigma^{K\mu} = \begin{pmatrix} A & & & \\ & B & & \\ & & C & 0 \\ & & 0 & C \end{pmatrix} \tag{29}$$

where K stands for A, B and C and we've left blank the region off of the three diagonal blocks. The elements of the above matrix are themselves 2×2 matrices so A is proportional to the 2×2 unit matrix.

The size of the Geo3 group is 6, so that the sizes of the three classes, given in the second line of the character table, sum to give the group size: $1+2+3=6$. Diagonalizing preserves the degrees of freedom so our diagonalization amounts to writing 6 as a sum of the squares of the sizes of the diagonal blocks:

$$6 = 1^2 + 1^2 + 2^2. \tag{30}$$

This shows how the 6 dimensions of $3\left[\sigma^{\mu}\right]$ as a vector space over the Pauli algebra appear in block diagonalized form.

Since the size of Geo3 is six, the $3\left[\sigma^{\mu}\right]$ algebra has enough degrees of freedom for six Weyl equations. Two equations are given by the A and B irreps. We can read them off of the A and B horizontal lines of the character table. The A irrep has character 1 for all the classes so all 6 group elements contribute equally and its Pauli spin algebra basis is given by summing over them:

$$\sigma^{A\mu} = \left(\sigma^{\mu(\,)} + \sigma^{\mu(123)} + \sigma^{\mu(132)} + \sigma^{\mu(12)} + \sigma^{\mu(13)} + \sigma^{\mu(23)}\right)/6. \tag{31}$$

The division by 6 is a normalization so that they square to their 2 × 2 identity $\sigma^{At} = I_2^A$.

The B irrep differs from the A irrep by having −1 for the {(12)} class (which are the three odd permutations), so it takes "−" signs for those elements:

$$\sigma^{B\mu} = \left(\sigma^{\mu(\,)} + \sigma^{\mu(123)} + \sigma^{\mu(132)} - \sigma^{\mu(12)} - \sigma^{\mu(13)} - \sigma^{\mu(23)}\right)/6. \tag{32}$$

Since the A and B Pauli spin algebra bases are in different diagonal blocks, they must annihilate each other as a matter of matrix multiplication. For example $\sigma^{A0}\sigma^{B3} = I_2^A\sigma_z^B = 0$. This is in contrast to the original basis where all the basis products are nonzero.

The 2 × 2 diagonal block will give four Weyl equations. It would be natural to label them according to their position in the 2 × 2 block. With this method, the four uncoupled Weyl equations would carry superscripts of $11,12,21,22$ and their corresponding Pauli spin algebras would fit into the 2 × 2 block as follows:

$$\sigma^{C\mu} = \begin{pmatrix} \sigma^{C11\mu} & \sigma^{C12\mu} \\ \sigma^{C21\mu} & \sigma^{C22\mu} \end{pmatrix}. \tag{33}$$

However, we'll be considering SU(2) transformations on these Equations (which correspond to SU(2) transformations on the internal symmetries of the corresponding particles) and our transformations will be simpler if we algebraically convert these four uncoupled Weyl equations labeled by $\{11,12,21,22\}$ into four new uncoupled Weyl equations with Pauli $\{t,x,y,z\}$ labels. With our usual notation abuse:

$$\begin{pmatrix} \sigma^{C11\mu} & \sigma^{C12\mu} \\ \sigma^{C21\mu} & \sigma^{C22\mu} \end{pmatrix} = \begin{pmatrix} \sigma^{Ct\mu} + \sigma^{Cz\mu} & \sigma^{Cx\mu} + i\sigma^{Cy\mu} \\ \sigma^{Cx\mu} - i\sigma^{Cy\mu} & \sigma^{Ct\mu} - \sigma^{Cz\mu} \end{pmatrix}. \tag{34}$$

So that, for example, $\sigma^{Ct\mu} = \left(\sigma^{C11\mu} + \sigma^{C22\mu}\right)/2$ and $\sigma^{Cz\mu} = \left(\sigma^{C11\mu} - \sigma^{C22\mu}\right)/2$. In the new notation, the basis element σ^{Cxy} is σ_x in the internal SU(2) and σ_y in the external or spin-1/2 SU(2) for the C irrep. The four equations $\left\{\sigma^{Ct\mu}, \sigma^{Cx\mu}, \sigma^{Cy\mu}, \sigma^{Cz\mu}\right\}$ can be manipulated by a transformation, for example,

$$\sigma^{Cx\mu} \to \sigma^{C'x\mu} = U\sigma^{Cx\mu}U^{-1} \tag{35}$$

where U is in SU(2). Transforming the generalized Pauli spin matrices using their internal SU(2) this way transforms the decoupled Weyl equations, that they define, to an equivalent set. We propose that this is a type of gauge transformation.

As an example of this internal SU(2) symmetry of the uncoupled Weyl equations, given an angle $\theta/2$, note that $(\theta/2)\sigma_x$ is in su(2) so that $U = \exp\left(i(\theta/2)\sigma_x\right)$ is in SU(2). This U leaves $\sigma^{Ct\mu}$ and $\sigma^{Cx\mu}$ unchanged but rotates the other two according to

$$\begin{aligned} \sigma^{C'y\mu} &= \cos(\theta)\sigma^{Cy\mu} + \sin(\theta)\sigma^{Cz\mu}, \\ \sigma^{C'z\mu} &= \cos(\theta)\sigma^{Cz\mu} - \sin(\theta)\sigma^{Cy\mu}, \end{aligned} \tag{36}$$

so that, as expected, U rotates the y and z components of $\sigma^{C\mu}$ by an angle θ. This simplifies our understanding of these four Weyl equations as an internal SU(2) symmetry so rather than using the natural matrix description of the four equations shown in Equation (33), we will use the Pauli algebra description of these four degrees of freedom which we can abbreviate as:

$$\sigma^{C\nu\mu} = \begin{pmatrix} \sigma^{Ct\mu} + \sigma^{Cz\mu}, \sigma^{Cx\mu} + i\sigma^{Cy\mu} \\ \sigma^{Cx\mu} - i\sigma^{Cy\mu}, \sigma^{Ct\mu} - \sigma^{Cz\mu} \end{pmatrix} \tag{37}$$

where $\nu \in \{t, x, y, z\}$ is the index for the Pauli algebra of the internal SU(2) symmetry, while μ is the usual index for the Pauli spin-1/2.

Our next task is to write out $\sigma^{C\nu\mu}$ in terms of the algebra. The first, $\sigma^{Ct\mu}$, is easiest; similar to $\sigma^{A\mu}$ and $\sigma^{B\mu}$ defined in Equations (31) and (32), it's given by the third line of the character table "$C \mid 2\ -1\ 0$ " and is:

$$\sigma^{Ct\mu} = 2\left(2\sigma^{\mu(\,)} - \sigma^{\mu(123)} - \sigma^{\mu(132)}\right)\Big/6. \tag{38}$$

The overall multiplication by 2 is needed to make it a projection operator and comes from the fact that the C irrep has a character of 2 for the identity.

To find the remaining three degrees of freedom, *i.e.* $\{\sigma^{Cx\mu}, \sigma^{Cy\mu}, \sigma^{Cz\mu}\}$, let us first reexamine the character table Equation (28). The three columns correspond to the three classes. These three degrees of freedom are rewritten as the three irreducible representations $\{A, B, C\}$. The three classes include all six of the group elements so we can rewrite this table from 3 × 3 to 6 × 3 if we replace the classes with their elements. For example, the second class {(123)} has two elements (123) and (132) so we expand the second column of the character table to two identical columns, the second and third of the expanded 6×3 character table:

Elements :	( )	(123)	(132)	(12)	(13)	(23)
A	1	+1	+1	+1	+1	+1
B	1	+1	+1	−1	−1	−1
C	2	−1	−1	0	0	0

(39)

To fill the table to 6 × 6 we need three more lines.

The three missing lines in the generalized character table correspond to the three internal SU(2) spin matrices $\sigma^{Cx\mu}$, $\sigma^{Cy\mu}$ and $\sigma^{Cz\mu}$. These lines have to be orthogonal to the three irreps we already have and they need to be orthogonal to each other. From the table we see that the three degrees of freedom left can be taken as $(123)-(132)$, $(12)-(13)$ and $(12)-(23)$. We could simply put these three lines into the table but they wouldn't quite act like Pauli spin matrices. They do span the internal Pauli spin matrices so we can write, for example:

$$(123)-(132) = \alpha\sigma_x + \beta\sigma_y + \gamma\sigma_z, \tag{40}$$

where α, β and γ are three complex numbers. Squaring both sides, and using the fact that the Pauli spin matrices anticommute we find

$$\begin{aligned}&\left[(123)-(132)\right]^2=\left[\alpha\sigma_x+\beta\sigma_y+\gamma\sigma_z\right]^2,\\&(132)+(123)-2(\ )=\left(\alpha^2+\beta^2+\gamma^2\right)\sigma_1=\left(\alpha^2+\beta^2+\gamma^2\right)\sigma^{Ct\mu}.\end{aligned}\tag{41}$$

The left hand side is the negative of the C irrep line. And since $\alpha^2+\beta^2+\gamma^2$ is just a complex number, we can take its square root and use it as a normalization factor on $(123)-(132)$ to give us $\sigma^{Cx\mu}$. We have freedom in our choice of α, β and γ and of course this freedom is precisely the internal SU(2) gauge freedom. We will choose $\beta=\gamma=0$ so that $\sigma^{Cx\mu}$ is a multiple of $(123)-(132)$. The multiple is determined by requiring that $\sigma^{Cx\mu}$ squares to $\sigma^{Ct\mu}$ given above as Equation (38).

Our choice of $\sigma^{Cx\mu}$ has used up the $(123)-(132)$ degree of freedom so the remaining two Pauli spin matrices, $\sigma^{Cy\mu}$ and $\sigma^{Cz\mu}$ will depend only on $(12)-(13)$ and $(12)-(23)$. Again the choice is arbitrary and corresponds to the su(2) gauge freedom that remains after defining σ_x. We will define $\sigma^{Cy\mu}$ as a multiple of $(12)-(13)$ and determine the multiplication coefficient by again requiring that it square to give $\sigma^{Ct\mu}$. Then we can find $\sigma^{Cz\mu}$ by $\sigma_x\sigma_y=i\sigma_z$. Including these new irreps in the expanded character table gives:

Elements :	$(\ )$	(123)	(132)	(12)	(13)	(23)
$\sigma^{A\mu}$	$1/6$	$+1/6$	$+1/6$	$+1/6$	$+1/6$	$+1/6$
$\sigma^{B\mu}$	$1/6$	$+1/6$	$+1/6$	$-1/6$	$-1/6$	$-1/6$
$\sigma^{Ct\mu}$	$2/3$	$-1/3$	$-1/3$	0	0	0
$\sigma^{Cx\mu}$	0	$i\sqrt{1/3}$	$-i\sqrt{1/3}$	0	0	0
$\sigma^{Cy\mu}$	0	0	0	$\sqrt{1/3}$	$-\sqrt{1/3}$	0
$\sigma^{Cz\mu}$	0	0	0	$1/3$	$1/3$	$-2/3$

(42)

This is a general result that does not depend on the choice of field or ring, so the μ was never used or needed. The squared magnitude of each line is the same as the coefficient for the identity, *i.e.* 1/6 for *A* and *B* and 2/3 for the four components of *C*. Each line of this table defines four transformation equations depending on the choice of μ. For example, the fourth equation of the last line is $\sigma^{Czz}=\sigma^{(12)z}/3+\sigma^{(13)z}/3-2\sigma^{(23)z}/3$. Note that the internal SU(2) degrees of freedom do not use the identity ( ) as we expect for traceless su(2) generators. The reader is invited to verify that they do form a representation of su(2).

Our new table upgrades the character table to a transformation on all the basis elements of the group algebra, and therefore, it is also a transformation on the group algebra itself. For Abelian groups of the form $G=\left\{a^k\mid 0\le k<N\right\}$ for some $a\in G$, the character table is of size $N\times N$ and is a discrete Fourier transform. Thus our new table is a generalization of the discrete Fourier transform to a non Abelian symmetry group.

In 2010 this author published a paper that rewrote the Koide equation for the masses of the charged leptons as the result of Feynman paths taken over spin in the *x*, *y* and *z* directions. [10] Given a path of *n* steps, the $n+1$ step will either be in the same direction, clockwise or counterclockwise around the $(1,1,1)$ direction. Three of the elements of Geo3 correspond to these choices: $(\ )=XYZ$,

$(123)=YZX$ and $(132)=ZXY$. They form the point group $\mathrm{Geo}\bar{3}$ which is an Abelian subgroup of Geo3 of size three so they imply a discrete Fourier transform on three objects as was pointed out by Marni Sheppeard in 2009. She asked "Is there a noncommutative transform that extends this analysis to nonclassical underlying spaces?" [11] This paper suggests that there is and that it will help understand the masses and symmetries of the Standard Model fermions. Rewriting the Koide equation allowed it to be extended to the neutrinos and, for the charged leptons, revealed another mass coincidence, a phase factor of 2/9. In 2012 Zenczykowski extended the 2/9 coincidence to the up and down quarks which take 1/3 and 2/3 of the charged lepton value [12]. This paper is another step in fully understanding the Standard Model fermions both symmetry and generation structure.

A Weyl equation is a matrix equation, so when we transform it according to its spin-1/2 SU(2) symmetry, we transform it the way we would a density matrix. That is, its Pauli matrices transform as $\sigma_j' = U\sigma_j U^{-1}$ like density matrices ρ do rather than by the way that state vectors transform $\psi' = U\psi$. Our treatment of the internal SU(2) symmetry of the C irreps as shown in Equation (35), uses the density matrix type transformation. Density matrices are convenient for calculations in quantum statistical mechanics so we briefly discuss them here.

The Gibbs form for the density matrix of an ensemble is

$$\rho(T) \propto \mathrm{e}^{-\frac{H}{T}}, \tag{43}$$

where the Hamiltonian H specifies the energy of the possible states, T is the temperature, the Boltzmann constant is unity, and we've left off the normalization factor that gives $\rho(T)$ a trace of 1. Squaring both sides gives

$$\rho^2(T) \propto \mathrm{e}^{-\frac{2H}{T}} = \mathrm{e}^{-\frac{H}{T/2}}, \tag{44}$$

so squaring such a density matrix gives a density matrix that is proportional to one with half the temperature. Repeating the procedure of squaring and dividing by the trace allows one to find a density matrix for a temperature close to zero. If the Hamiltonian depends on the particle states, this low temperature limit will be the pure density matrix corresponding to the state with the lowest energy.

The high temperature limit for a density matrix is given by letting $T \to \infty$ and is proportional to the unit matrix. The proportionality constant is the trace so the high temperature limit for $3\left[\sigma^{\mu}\right]$ is

$$\rho(T=\infty) = (\ )/6, \tag{45}$$

as the trace of the identity of a group algebra is the size of the group (and the other basis elements of the group algebra have zero trace). Density matrices have to be Hermitian and the same principle applies to a group algebra. The generalization of "Hermiticity" from a matrix to a group algebra is covered in introductory graduate math texts or the reader can work backwards from the Hermiticity requirement for the matrix form of the algebra.

If we choose a small random Hermitian element of the group algebra we can

add it to the high temperature limit to get a Hermitian element that is a valid state near the high temperature limit. Cooling this state down to a nearly pure state gives us a random pure state. For $3\left[\sigma^{\mu}\right]$ there are three possible pure states, A, B and C. Figure 1 shows the result of this cooling for a few thousand random high temperature states. The A and B irreps appear as two dots on the right side of the images while the internal SU(2) C state appears as a Bloch sphere on the left. The algebra has six group degrees of freedom and the images are two dimensional so four degrees of freedom are not graphed. The degrees of freedom that are graphed have been selected to separate the A, B and C states and to spread the C states with $\sigma^{Cx\mu}$ in the x-direction and $\sigma^{Cy\mu}$ in the y-direction so the Bloch sphere appears as an x-ray of a sphere. If you assign a graduate student to read and understand this paper, a suitable task is to reproduce this graph for this or other point group symmetries.

In this paper the ring we use the Pauli algebra but the figure does not depend on the field so when writing a computer program to plot the cooling process it is more convenient to use the complex numbers as the field. Then an element of the group algebra $3[\mathbb{C}]$ consists of six complex numbers. Addition is simply vector addition. To multiply algebra elements the computer program uses the Geo3 group multiplication rules. Repeatedly squaring and dividing by the trace creates a path that is then graphed.

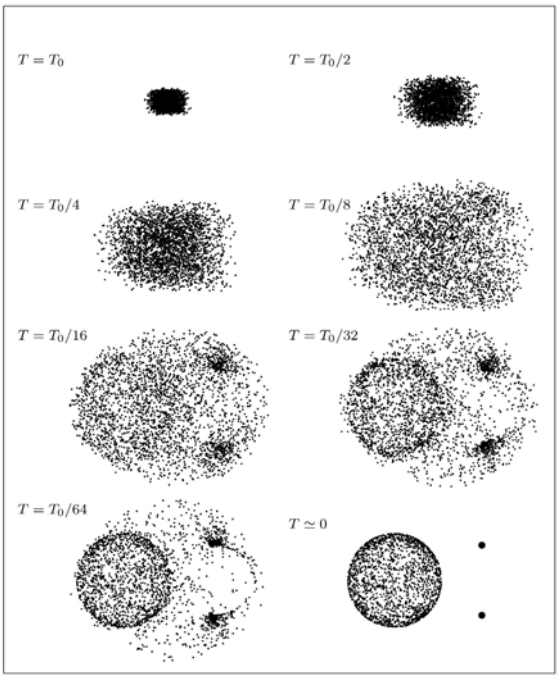


Figure 1. Beginning with 3000 random density matrices near the high temperature limit $T = T_0$, we square the density matrix six times to show the states beginning to converge to the two singletons A and B, and the internal SU(2) doublet C. Of the six degrees of freedom in the $3\left[\sigma^{\mu}\right]$ algebra, we choose the x and y axes so as to spread A, B and C apart, and to show the Bloch sphere for C. Continuing the cooling process, the final image $T \simeq 0$ shows the cold temperature limit (pure states).

4. $43\left[\sigma^{\mu}\right]$ and the Standard Model

In this section we will be discussing a single generation of the Standard Model, *i.e.* the electron, neutrino, up quark and down quark. In the usual quantum mechanical description of these particles, each is modeled using bispinors and the Dirac equation. Bispinors have four complex components. For example the electron bispinor basis can be chosen to be the four states "spin-up electron", "spin-down electron", "spin-up positron" and "spin-down positron". An alternative natural to the Weyl basis would be "spin-up left", "spin-down left", "spin-up right" and "spin-down right". We cannot define a basis of elements that mix handedness with charge, *i.e.* "spin-up left positron" in the sense of a particle whose measured electric charge is +1 because charge and handedness do not commute. However, if we treated the positron as an electron travelling backwards in time, then the "unphysical" electric charge, that assigns −1 to both the electron and the positron, does commute with handedness.

We will avoid the mass interaction in this section. In the massless limit (relativistic or high energy) the Dirac equation splits into two uncoupled Weyl equations, left handed and right handed. Each of these equations is a mix of particle and antiparticle. We will be using the quantum information limit which amounts to an infinite mass. In this limit the Dirac equation also splits into two Weyl type equations, but massive, one is for particles the other for antiparticles and the momentum is required to be zero. The $43\left[\sigma^{\mu}\right]$ algebra includes both left handed and right handed coupled Weyl equations, but the irreducible representations will mix these leaving particles and antiparticles. So our uncoupled Weyl equations will have a basis of "spin-up electron", "spin-down electron", "spin-up positron" and "spin-down positron". This makes it easy to find the operator for electric charge. By looking at the left and right handed parts of the algebra, we can pick out the left and right handed parts of the particles and therefore we can derive operators for weak hypercharge and weak isospin.

An alternative to our calculation would be to write a generalized massless Dirac equation using the handed octahedral point group $\mathrm{Geo}\overline{43}$, which has 24 elements and is the subgroup consisting of Geo43's proper rotations. Then the algebra would be $\overline{43}\left[\gamma^{\mu}\right]$. The reason we've chosen $43\left[\sigma^{\mu}\right]$ instead is that it makes the projection operators for handedness part of the point group symmetry. With $\overline{43}\left[\gamma^{\mu}\right]$, the handedness projection operators are $\left(1\pm\gamma^{5}\right)/2$ which would mean our particle identities would be defined by a combination of the point group symmetry and the gamma matrices.

Half of the 24 elements of the Geo43 point group are proper rotations and leave the Pauli spin matrices left handed. The other 24 elements change the handedness of the Pauli spin matrices to right handed. The irreducible representations exist in two types. Half have the same signs for the left-handed and right handed elements, the other have opposite signs. The usual convention is that particles take the representations with the same signs while the antiparticles use the representations with opposite signs. In addition to the

difference in handedness, the basis elements of Geo43 can also be classified as "even" and "odd". The proper rotations which can be obtained as an even number of right angle rotations, and the improper rotations which are a proper even rotation times the inversion i are the "even" elements, the rest are "odd". For the left handed particles, weak hypercharge depends on left handed even elements while weak isospin depends on left handed odd elements.

It's clear that the lepton irreps will correspond to 1×1 blocks on the diagonal while the quarks will be 3×3. With these assumptions, it remains to distinguish between the two leptons, *i.e.* electron and neutrino, and between the two quarks. Until we have a more complete theory that includes the gauge bosons, we will make an arbitrary choice for these. Labeling the irreps of the character table with the particles and antiparticles we have two irreps left over that we assign to dark matter D and its antiparticle $\overline{D}$:

O	E	$3C_2$	$8C_3$	$6C_2$	$6C_4$	i	$3\sigma_h$	$8S_6$	$6S_4$	$6\sigma_d$
Hand:	L	L	L	L	L	R	R	R	R	R
Ev/Od:	E	E	E	O	O	E	E	E	O	O
Size:	1	3	8	6	6	1	3	8	6	6
σ^μ:	XYZ	$X\overline{YZ}$	YZX	$Y\overline{XZ}$	$X\overline{ZY}$	$\overline{XYZ}$	$\overline{XYZ}$	$\overline{YZX}$	$\overline{YXZ}$	$\overline{XZY}$
ν	1	1	1	1	1	1	1	1	1	1
$\overline{\nu}$	1	1	1	1	1	−1	−1	−1	−1	−1
e	1	1	1	−1	−1	1	1	1	−1	−1
$\overline{e}$	1	1	1	−1	−1	−1	−1	−1	1	1
d	3	−1	0	1	−1	3	−1	0	1	−1
$\overline{d}$	3	−1	0	1	−1	−3	1	0	−1	1
u	3	−1	0	−1	1	3	−1	0	−1	1
$\overline{u}$	3	−1	0	−1	1	−3	1	0	1	−1
D	2	2	−1	0	0	2	2	−1	0	0
$\overline{D}$	2	2	−1	0	0	−2	−2	1	0	0

(46)

In the above table, the five columns on the left are the left handed or proper rotations, and the five columns on the right are the right handed or improper rotations. Each of these five columns are split into three even and two odd columns.

We can write the operator for electric charge Q in terms of the classes of the Geo43 group. That is, Q is a diagonal operator in $43\left[\sigma^\mu\right]$ that commutes with all the irreps and so can be written as a sum

$$Q = \Sigma_j q_j I_j \tag{47}$$

where q_j is the electric charge of the jth particle, and I_j is the identity for the corresponding block on the diagonal. The blocks on the diagonal correspond to the irreps of the original character table and so each I_j can be written as a sum of products of the group classes. These can be read off of the character table as shown in the previous section. As always occurs with character tables, the number of irreps is the same as the number of classes (columns) so there is only one choice for the electric charge operator when writing it as an element of $43\left[\sigma^\mu\right]$.

Since particles and anti particles have opposite electric charge, examining the table shows that the Q operator can be composed only of right handed elements. The left (right) handed particles will be composed only of left (right) handed elements so their weak isospin and weak hypercharge will depend only on left

(right) handed group elements. Weak hypercharge for right handed particles is the same as electric charge which we already know is made only of right handed elements. This reduces the algebra needed to define the charges to be a 5×5 transformation instead of 10×10. Writing the left and right handed elements of the particles in the same columns we have:

$L:$	E	$3C_2$	$8C_3$	$6C_2$	$6C_4$				
$R:$	i	$3\sigma_h$	$8S_6$	$6S_4$	$6\sigma_d$				
Hand:	L/R	L/R	L/R	L/R	L/R				
Ev/Od:	E	E	E	O	O	Charge			
Size:	1	3	8	6	6	$Y/2$	$+ I_3$	$= Q$	
ν	1	1	1	1	1	$\frac{-1}{2}/0$	$\frac{1}{2}/0$	0	
e	1	1	1	-1	-1	$\frac{-1}{2}/-1$	$\frac{-1}{2}/0$	-1	
d	3	-1	0	1	-1	$\frac{1}{6}/\frac{-1}{3}$	$\frac{-1}{2}/0$	$\frac{-1}{3}$	
u	3	-1	0	-1	1	$\frac{1}{6}/\frac{2}{3}$	$\frac{1}{2}/0$	$\frac{2}{3}$	
D	2	2	-1	0	0	0/0	0/0	0.	

(48)

When the classes in the above table are expanded to the elements the five columns will increase to $1+3+8+6+6=24$ and the operators are obtained by dividing by 24. We will do the algebra without this division and insert it at the end. Also, in the above table there is a subtlety on the scaling. The projection operators for the quark identities 1_3 is three times the amount shown. This would require us to multiply those rows by three, but these projection operators apply to three colors of quarks while the experimentally measured charges are for a single quark so we also need to divide by three.

Right weak isospin is always zero, and charge is the same as right weak hypercharge leaving us three equations to solve: left and right weak hypercharge, and left weak isospin. Left weak hypercharge depends only on the even left components giving three equations in three unknowns:

$$\begin{aligned} &1Y/2^{E} + 1Y/2^{3C2} + 1Y/2^{8C3} = -1/2, \\ &3Y/2^{E} - 1Y/2^{3C2} + 0Y/2^{8C3} = 1/6, \\ &2Y/2^{E} + 2Y/2^{3C2} - 1Y/2^{8C3} = 0. \end{aligned} \tag{49}$$

Left weak isospin depends only on the odd left components so there are two equations in two unknowns:

$$\begin{aligned} &1I_3^{6C2} + 1I_3^{6C4} = 1/2, \\ &1I_3^{6C2} - 1I_3^{6C4} = -1/2. \end{aligned} \tag{50}$$

Electric charge and right weak hypercharge use all five right components giving five equations:

$$\begin{aligned} &1Y/2^{i} + 1Y/2^{3\sigma h} + 1Y/2^{8S6} + 1Y/2^{6S4} + 1Y/2^{6\sigma_d} = 0, \\ &1Y/2^{i} + 1Y/2^{3\sigma h} + 1Y/2^{8S6} - 1Y/2^{6S4} - 1Y/2^{6\sigma_d} = -1, \\ &3Y/2^{i} - 1Y/2^{3\sigma h} + 0Y/2^{8S6} + 1Y/2^{6S4} - 1Y/2^{6\sigma_d} = -1/3, \\ &3Y/2^{i} - 1Y/2^{3\sigma h} + 0Y/2^{8S6} - 1Y/2^{6S4} + 1Y/2^{6\sigma_d} = 2/3, \\ &2Y/2^{i} + 2Y/2^{3\sigma h} - 1Y/2^{8S6} + 0Y/2^{6S4} + 0Y/2^{6\sigma_d} = 0. \end{aligned} \tag{51}$$

Solving these linear equations for the operator coefficients defines the operators as

$$\begin{array}{rcccccc} I_3 & = & 0 & 0 & 0 & 0 & +12(6C_4), \quad \text{Left} \\ Y/2 & = & 0 & -4(3C_2) & -8(8C_2) & 0 & 0 \quad \text{Left} \\ & & 0 & -4(3\sigma_h) & -8(8S_6) & 0 & +12(6\sigma_d), \quad \text{Right} \\ Q & = & 0 & -4(3\sigma_h) & -8(8S_6) & 0 & +12(6\sigma_d), \quad \text{Right} \end{array} \tag{52}$$

where the terms have been arranged according to their symmetry and their left and right elements are labeled to the right side. The coefficients are for the class; we could have divided by the size of each class and obtained charges per group element of −4/3 for $(3C_2)$ and $(3\sigma_h)$, −1 for $(8C_2)$ and $(8S_6)$, and +2 for $(6C_4)$ and $(6\sigma_d)$.

In mapping the irreps to the particles we had freedom in that we could independently swap the leptons and swap the quarks. This would change some of the signs in the −4 and −8 columns as well as swapping $(6C_4)$ for $(6C_2)$ and $(6S_4)$ for $(6\sigma_d)$.

Since charge does not commute with handedness, the equation $Y/2 + I_3 = Q$ requires an assumption not obvious in the above: it follows when one uses Q as the charge of a Dirac bispinor so that the particle portion has charge Q while the antiparticle portion has charge $-Q$.

For low temperatures, a heat bath is typically conceived as an object that can exchange electric field gauge bosons (*i.e.* heat photons) with the system under study. However, when we model such a system with density matrices the gauge bosons do not appear explicitly; their effect is implicit. The gauge boson for the weak force is the $W^{\pm}$. As the temperature rises above the amount needed to create $W^{\pm}$ gauge bosons, it becomes possible for the heat bath to convert between electrons and neutrinos, and to convert between up quarks and down quarks. In the temperature flow shown in **Figure 1**, we can see that the cooling states clump together into curving arcs. These occur because certain degrees of freedom cool faster than others. We can imagine that these correspond to a cascade of symmetry breaking from the high temperature limit where all states are approximately identical. Then the weak symmetry breaking is the coolest of these.

5. Generations

While the group algebra Geo43 discussed in the previous section includes a single generation of the Standard Model fermions, to include the generation structure requires that we use a finite group that has three times as many degrees of freedom. The finite Abelian group of size 3, C_3, is perfect for this. To define this group we will use the complex cubed root of unity

$$\omega = \exp(2i\pi/3). \tag{53}$$

So the group elements are $\{1, \omega, \omega^*\}$ and the group product is the usual for complex numbers. The character table for Z_3 is:

C_3	1	ω	ω^*
A	1	1	1
E	1	ω	ω^*
E^*	1	ω^*	ω

(54)

Since this is an Abelian group, the three classes correspond to the three elements of the group and the character table defines a discrete Fourier transform.

When we multiply a finite group by an Abelian finite group the character table of the new group is the product of the two character tables. Thus when we triple Geo43 with C_3 to form the modified point group we call 3Geo43 it will have 30 classes instead of the 10 that Geo43 had, and it will have 30 irreps as well. Each of the old Geo43 irreps will appear in the new character table three times, once with $(1,1,1)$ as the characters of $(1,\omega,\omega^*)$, again with characters $(1,\omega,\omega^*)$ and finally with characters $(1,\omega^*,\omega)$. Thus ν will now appear three times, once with thirty 1s, once with ten 1s, ten ωs and ten ω^*s, and finally the same but with ω and ω^* reversed.

The character table of 3Geo43 will be of size 30 × 30 and is larger than we can reasonably print in a journal. There will now be 144 Weyl equations to decouple with a total of $2\times 4\times 144^2 = 165888$ terms but the equations will uncouple according to the character table as before and the result will be three copies or generations of the Standard Model fermions.

Given that the character table of C_3 defines the generation structure, it's natural to look for a dependency on those characters in experimental data that depends on generation. The most obvious of these are how the masses of the Standard Model fermions depend on the generation and indeed, this is precisely what is seen in the Koide formulas for the lepton and quark masses [12] [13] [14]. Other generation structures, such as the CKM and PMNS matrices that define how the weak force converts between generations will be more complicated in that they must depend on both the C_3 and Geo43 character tables and formulas for those have not yet been found.

6. Dark Matter Nomenclature

The extra "dark matter" irreps, D and $\bar{D}$ in the $43\left[\sigma^{\mu}\right]$ character table of Equation (46) have zero electric charge, weak hypercharge and weak isospin and so do not participate in the electric or weak forces. In the Standard Model, the color force is determined by quark SU(3) internal symmetries. In this model, the quark SU(3) operators are internal to the quark irreps in that they act on the traceless degrees of freedom in the 3 × 3 quark blocks. The remaining degree of freedom is the 3 × 3 unit matrix which is the projection operator proportional to the usual irrep. These quark SU(3) color degrees of freedom are restricted to the corresponding quark 3 × 3 blocks so they annihilate between different quarks.

Examining the character table we see that the quarks differ in the signs of their odd elements, $(6C_2)$, $(6C_4)$, $(6S_4)$ and $(6\sigma_d)$ therefore the gluons must be associated with changes in these elements. But the dark matter irreps are zero in those columns so they cannot participate in the strong force. Thus the dark matter irreps are indeed dark.

As with the leptons and quarks, dark matter comes in particle D and

antiparticle $\bar{D}$ form and appears in three generations. They differ in sign between the left and right handed parts of the particles as do the leptons and quarks so we expect that dark matter will also have mass.

The dark matter irreps have an internal SU(2) symmetry that is otherwise similar to the quark's internal SU(3) symmetry. By analogy with the quarks, we will label the dark matter SU(2) internal symmetry as colors. As the dark matter particles that correspond to quarks, we will call them "dark quarks" or "duarks". Since these have only an SU(2) symmetry, there are only two dark colors needed as a basis for dark color. Where the colors for quarks are expressive of their participation in electromagnetic photon interactions, the duarks are dark so in contrast to the quark colors of red, green and blue, we will use dark colors of "doom" and "gloom". These are also appropriate for this paper which is being written under the looming threat of the Covid19 pandemic.

We assume that there is a force boson similar to the gluon between them, call it the "duon", that will follow an SU(2) triplet symmetry. Where the eight gluons are often described using a basis of the eight Gell-Mann matrices, the three duons can use the three Pauli matrices to define the dark matter triplet:

$$\begin{aligned} &\left(d\bar{g}+g\bar{d}\right)/\sqrt{2}, \\ &-i\left(d\bar{g}-g\bar{d}\right)/\sqrt{2}, \\ &\left(d\bar{d}-g\bar{g}\right)/\sqrt{2} \end{aligned} \tag{55}$$

where d and g stand for the doom and gloom dark colors. The presence of a gluon-like force implies that dark matter has a significant scattering cross section with itself. This possibility is in the literature described as "self interacting cold dark matter" or SICDM [14].

A quark and an antiquark can combine to form a meson; we expect that a duark and an antiduark similarly combine to form a "deson". While the mesons decay by electroweak processes, the desons do not participate so we may suppose that they are stable. Similarly, three quarks can combine to form a baryon so two duarks of different dark colors, combine to form a "daryon".

Acknowledgements

We hope that the reader has enjoyed reading this paper even more than the author enjoyed writing it. Of course the author has received assistance from many physicists. Of particular note are his advisors in the physics program at Washington State University, Michael Forbes, Sukanta Bose and Fred Gittes. Their interest and encouragement were critical to maintaining the long effort required here. And that long effort was too long by two or three years as WSU has a limitation on how long one can take to write a thesis. The author thanks his advisors for continuing to assist after that time ran out. The author hopes that the university will find a way of making an exception and that this paper will be considered as a partial fulfilment of the PhD degree in physics at WSU.

Conflicts of Interest

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

References

[1] Peskin, M.E. and Schroeder, D.V. (1995) Introduction to Quantum Field Theory. Addison-Wesley Pub. Co.

[2] Hestenes, D. and Holt, J.W. (2007) *Journal of Mathematical Physics*, **48**, Article ID: 023514. https://doi.org/10.1063/1.2426416

[3] Terras, A. (2012) Harmonic Analysis on Symmetric Spaces and Applications I. Springer Science & Business Media, Berlin.

[4] Mackey, G.W. (1980) *Bulletin of the American Mathematical Society*, **3**, 543-698. https://doi.org/10.1090/S0273-0979-1980-14783-7

[5] Steinberg, B. (2012) Representation Theory of Finite Groups: An Introductory Approach. Springer, Berlin. https://doi.org/10.1007/978-1-4614-0776-8

[6] Patrignani, C., *et al.* (2016) *Chinese Physics C*, **40**, Article ID: 100001. https://doi.org/10.1088/1674-1137/40/10/100001

[7] Roe, B.P. (1996) Particle Physics at the New Millennium. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4612-2362-7

[8] Dreiner, H.K., Haber, H.E. and Martin, S.P. (2010) *Physics Reports*, **494**, 1-196. https://doi.org/10.1016/j.physrep.2010.05.002

[9] Hamermesh, M. (1962) Group Theory and Its Application to Physical Problems. Dover, Mineola. https://doi.org/10.1119/1.1941790

[10] Brannen, C.A. (2010) *Foundations of Physics*, **40**, 1681-1699. https://doi.org/10.1007/s10701-010-9465-8

[11] Sheppeard, M. (2009) *Rejecta Mathematica*, **1**, 104-110.

[12] Zenczykowski, P. (2012) *Physical Review D*, **86**, Article ID: 117303. https://doi.org/10.1103/PhysRevD.86.117303

[13] Brannen, C.A. (2006) Koide's Mass Formula for Neutrinos. 8*th Annual APS Northwest Section Meeting*, Tacoma, 19-20 May 2006. http://meetings.aps.org/Meeting/NWS06/Event/50546

[14] Spergel, D.N. and Steinhardt, P.J. (2000) *Physical Review Letters*, **84**, 3760-3763. https://doi.org/10.1103/PhysRevLett.84.3760

Journal of Modern Physics, 2020, 11, 1215-1235
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Possible Solution to the Disagreement about the Hubble Constant II

Frank R. Tangherlini

San Diego, CA, USA
Email: frtan96@gmail.com

How to cite this paper: Tangherlini, F.R. (2020) A Possible Solution to the Disagreement about the Hubble Constant II. *Journal of Modern Physics*, **11**, 1215-1235.
https://doi.org/10.4236/jmp.2020.118076

Received: July 22, 2020
Accepted: August 18, 2020
Published: August 21, 2020

Abstract

This work continues the previous study (2018) *Journal of Modern Physics*. 9, 1827-1837, that proposes that the disagreement arises because the cosmic microwave background (CMB) value for the Hubble constant H_0 is actually for a universe which is decelerating rather than accelerating. It is shown that when H_0 of Freedman *et al.* (2019) *Astrophysical Journal*, 882: 34 (24 pp.) is re-determined for redshift $z = 0.07$, by replacing $q_0 = -0.53$ with $q_0 = 0.50$, the new lower value is in excellent agreement (0.1%) with the CMB H_0. The model is modified to include the clustering of galaxies, and the recognition that there are clusters that do not experience the Hubble expansion, such as the Local Group, and hence, in accordance with the model, within the Local Group the speed of light is c. The bearing of this result on the neutrino and light time delay from SN1987a is discussed. It is suggested that the possible emission of a neutrino from the blazar TXS-0506+56, that was flaring at the time, as well as possible neutrino emission earlier, may arise instead from a more distant source that happens to be, angle-wise, near the blazar, and hence the correlation is accidental. The model is further modified to allow for a variable index of refraction, and a comparison with the ΛCDM model is given. The age of the universe for different values of H_0 is studied, and comparison with the ages of the oldest stars in the Milky Way is discussed. Also, gravitational wave determination of H_0 is briefly discussed.

Keywords

Hubble Constant Disagreement, Decelerating Universe, Galactic Clusters, Age of Universe

1. Introduction

In the previous work [1] that dealt with the disagreement about the Hubble con-

stant H_0, it was proposed that the cosmic microwave background (CMB) [2] and the baryon acoustic oscillation [BAO] [3] values for H_0, although based on the flat ΛCDM model of an accelerating universe, with the dark energy possessing a negative pressure supplied by the cosmological term [4] [5] [6], could alternatively be interpreted as being for a decelerating universe, with the dark energy possessing rather an index of refraction, as described in previous works [7] [8] [9] [10]. It was pointed out that this alternative interpretation was possible because the CMB and closely matched BAO determinations do not involve the deceleration parameter q_0, and hence they are distance determinations of H_0, as was shown explicitly in [1] for the BAO determination. Consequently, if an alternative model based on a decelerating universe can get the same distances, to within experimental error, it follows that the CMB and BAO determinations can alternatively be thought as for a decelerating universe. Thus, as was noted in [1], *there is an ambiguity in the CMB and BAO determinations of* H_0. To resolve this ambiguity, it was noted that the SHoES local distance ladder (LDL) determination (Note: Formerly described as cosmic distance ladder (CDL) determination.) of H_0 of Riess *et al.* [11] does in fact make use of the deceleration parameter q_0, with $q_0 = -0.55$, and hence this determination *unambiguously* describes an accelerating universe. It was further noted that the empirical inequality $H_0(\mathrm{LDL}) > H_0(\mathrm{CMB})$ can be readily understood if the Hubble constant $H_0(\mathrm{CMB})$ is actually for a decelerating universe, rather than an accelerating universe, as is customarily assumed. This follows from the fact that since $H_0 = \dot{a}_0/a_0$, at the present epoch, where $a(t)$ is the expansion parameter for the Friedmann, LeMaître, Robertson, Walker (FLRW) line-element, and since $a_0(\mathrm{LDL}) = a_0(\mathrm{CMB})$, because the diminished brightness of the Type Ia supernovae (SNe Ia) has to be the same for both models, it follows $\dot{a}_0(\mathrm{LDL}) > \dot{a}_0(\mathrm{CMB})$. This latter inequality is what one expects if $H_0(\mathrm{CMB})$ is for a decelerating universe, assuming that for a sufficiently large redshift, $\dot{a}(t)$ was essentially the same for both models. Then, because of the acceleration of the one model, and the deceleration of the other, at the present epoch, $\dot{a}_0$ will be greater for the accelerating model than for the decelerating model, and hence the above inequality. Thus, in effect, *the decelerating model predicts the sign of the inequality between the two competing values of the Hubble constant.* It was also shown in [1] that if in the determination of $H_0(\mathrm{LDL})$ a positive deceleration parameter was employed rather than a negative one, $H_0(\mathrm{LDL})$ could be brought down towards $H_0(\mathrm{CMB})$, to within a correction involving the ratio of the distances for the two models for a given redshift, with the decelerating model requiring a greater distance. It was therefore suggested in [1] that astronomers should re-determine the distances in the LDL work, to see whether, when this was done, and the decelerating parameter changed from one for an accelerating universe to one for a decelerating universe, the resulting new value of H_0 would agree with $H_0(\mathrm{CMB})$. This suggestion was further described in a contributed talk [12] delivered at the American Physical Society April 2019 meeting, as will be discussed further below. A few months later, the author learned that a new

LDL determination of H_0 had been made by the Carnegie-Chicago collaboration of Friedman *et al.* [13] that employed the Tip of the Red Giant Branch (TRGB) for local distance determinations, rather than the Cepheids that were used in the SHoES team of Riess *et al.* [11]. Importantly, the Carnegie-Chicago work gave a lower value of H_0, corresponding to a greater distance than SHoES for a given redshift. This finding can be taken as support for the view expressed above that the redshift distances are greater for the proposed model than for SHoES; this will be discussed in greater detail in Section 2. In Section 3, the clustering of galaxies, that was not taken up previously, will be taken up, and its importance for the model described. This discussion will include a less *ad hoc* explanation for the relatively short arrival time difference between photons and neutrinos from SN1987a than what was given in [10]. Also, in Section 2, a possible objection to the model (not referenced in [1], but in [12]) will be taken up that is based on the 3σ correlation of gamma rays with a neutrino that was possibly emitted from a blazar that was flaring at the time [14]. *Importantly, because of the cosmological distances involved, such a correlation, if confirmed by later work, would falsify the proposed model.* In Section 4, the index of refraction n that was taken to be a constant in the model for the redshift range $0 \le z \le 1.0$ originally, and that was later in [10] restricted to the range $0 \le z \le 0.6$, will now be assumed to vary linearly for the range $0 \le z \le 1.7$, with $n = 1$ for $z \ge 1.7$. The choice of the rounded-off value $z = 1.7$ goes back to the finding of Riess *et al.* [15] that at $z = 1.65 \pm 0.15$ there was no acceleration. A comparison of the revised model with the ΛCDM is given in **Table 1**, and its implication is discussed. In Section 5, the age of the universe from the standpoint of the flat ΛCDM model, and the various values of H_0 associated with this model will be examined, and compared with the age of the oldest stars in the Milky Way. In Section 6, there are concluding remarks.

2. Possible Support for the Model from the TRGB Determination of H_0

As was pointed out in the Introduction, if the value found for $H_0(\mathrm{CMB})$ is interpreted as being for a decelerating universe, rather than for an accelerating universe, the proposed model predicts the sign of the inequality that has been found from observation to be $H_0(\mathrm{LDL}) > H_0(\mathrm{CMB})$. In order to clarify this interpretation further, assume: 1) that the decelerating model gives the correct distances as a function of redshift, as will be justified below from the excellent agreement that is obtained with $H_0(\mathrm{CMB})$, and 2) that the distances from the ΛCDM model provide a good approximation to the distances from the decelerating model that results in the ambiguity described above in the Introduction. Then, in obtaining the value of H_0 from the distances associated with the ΛCDM model, it would be quite reasonable to assume, as the Planck collaboration does, that one is determining the Hubble constant for an accelerating universe. This would also be the case for the BAO determinations. Importantly, it is possible to verify the above re-interpretation of $H_0(\mathrm{CMB})$ by re-determining the value

for $H_0(\mathrm{LDL})$ by replacing the negative deceleration parameter currently used, $q_0 = -0.55$, with the positive deceleration parameter, $q_0 = 0.50$, appropriate to the Einstein de Sitter (EdS) decelerating universe, to see whether the subsequently revised value of the Hubble constant comes down reasonably close to the lower CMB value. In Equation (10) in [1] it was shown that the logarithms for the ratio of the Hubble constants for the two different values of the deceleration parameter q_0 could be written as

$$\log\left(\frac{H_0(-0.55)}{H_0(0.50)}\right) = M - M' + \log\left(\frac{1+0.775z-0.274z^2}{1+0.25z-0.125z^2}\right), \tag{1}$$

where M and M' are the redshift-dependent distance moduli for the accelerating and decelerating universes, respectively, and the next term, in which z is the redshift, involves the kinematic corrections to the lowest order Hubble-LeMaître relation

$$k(z, q_0, j_0) = \left\{1+\frac{1}{2}\left[1-q_0\right]z-\frac{1}{6}\left[1-q_0-3q_0^2+j_0\right]z^2\right\}, \tag{2}$$

where j_0 is the jerk given by $\dddot{a}_0 a_0^2/\dot{a}_0^3$ [16] [17] [18] [19], and is taken to be unity in [11], while for the decelerating EdS universe, with $a(t) = Kt^{2/3}$, j_0 is also unity, independently of epoch. If, as was done in [12], one now introduces luminosity distances corresponding to the distance moduli in (10), defined by $-\log D_L \equiv M$, and $-\log D_L' \equiv M'$, and removes the logarithms in (1), and solves for $H_0(0.50)$, one obtains

$$H_0(0.50) = H_0(-0.55)\frac{D_L}{D_L'}\left(\frac{1+0.25z-0.125z^2}{1+0.775z-0.274z^2}\right). \tag{3}$$

Since there is going to be a comparison with the work in [13] further below, instead of the values for $H_0(-0.55)$ and z that were used in [1] and [12], the values used here will be $H_0(-0.55) = 74.03 \pm 1.42\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$, as given in [20], and in order to stay within the range of range of redshift values used in [13] that satisfied $z \le 0.08$, the value of redshift used in (3) for comparison will be $z = 0.07$ so that $k(0.07, 0.50, 1)/k(0.07, -0.55, 1) = 0.966$ and hence (3) yields

$$H_0 = (D_L/D_L')71.51 \pm 1.37\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}. \tag{4}$$

One sees that the SHoES value of $H_0(\mathrm{LDL})$ does come down towards the 2018 Planck CMB value of $H_0 = 67.36 \pm 0.54\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$ [21], but to fully attain it, it must be the case that the distance predicted by the model D_L' is greater than that given by SHoES. It was therefore suggested in [12] that the astronomical community re-determine the LDL distances to see if this would turn out to be the case. Quite surprisingly, almost as if to answer this suggestion, as remarked in the Introduction, there appeared the Carnegie-Chicago work of Freedman *et al.* [13], that found $H_0 = 69.8 \pm 0.8(\mathrm{stat}) \pm 1.7(\mathrm{sys})$, based on obtaining distances from the observation of the TRGB, that lends support to the model's requirement that the correct distances are larger than those obtained from the Cepheid measurements used in the SHoES determination of $H_0(\mathrm{LDL})$. To show that the TRGB finding does indeed lend support, account must be taken of the fact

that in [13] $q_0 = -0.53$ was used, since later determination led to $\Omega_m = 0.315 \pm 0.007$, that was simplified to $\Omega_m = 0.315$, rather than $\Omega_m = 0.308$ that was used in [11], and since $q_0 = \frac{1}{2}\Omega_m - \Omega_\Lambda$, this led to the above change in the q_0 used in [13]. Also, one must take into account that, since in [13] lower values of redshift were used than in [11], their form for the kinematic factor $k(z, q_0, j_0)$ omitted the z^2 term. However, if one omits it in obtaining (4), but keeps the other terms the same, one finds the ratio of the kinematic factors is still 0.966, so the omission of the z^2 term in the kinematic factor is not significant for this low value of redshift. Next, without changing the luminosity distances found in [13], so that $D_L' = D_L$, but changing only the deceleration parameter from $q_0 = -0.53$ to $q_0 = 0.50$, one has that (3) reduces to

$$H_0(0.50) = H_0(-0.53)\left(\frac{1+0.25z}{1+0.765z}\right). \tag{5}$$

Then, with $H_0(-0.53) = 69.8 \pm 2.0\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$ from [13], where the uncertainties have been rounded to ±2.0 $\mathrm{kms^{-1} \cdot Mpc^{-1}}$, and with $z = 0.07$, one has

$$H_0(0.50) = 67.43 \pm 1.93\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}, \tag{6}$$

in excellent agreement (0.1%) with the above CMB value [21]. However, *this agreement is for only one redshift*, and hence it was suggested to the Carnegie-Chicago collaboration that their value for H_0 be recalculated for the entire range or redshifts they used, upon replacing $q_0 = -0.53$, with $q_0 = 0.50$, in order to compare with the CMB value of H_0; but at this writing, this has not been done. The value in (6) is also in excellent agreement with the latest BAO results, as given in Addison *et al.* [22] for which $H_0 = 66.98 \pm 1.18\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$, Macaulay *et al.* [23], for which $H_0 = 67.8 \pm 1.3\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$, and that of Ryan *et al.* [24], for which $H_0 = 67.78^{+0.91}_{-0.87}\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$: Since, if one takes the weighted average of the three BAO measurements of H_0, with the fractional weights w_i given by $\sigma_i^{-2} / \sum_i \sigma_i^{-2}$, $i = 1, 2, 3$, one obtains

$$\langle H_0 \rangle = \sum_i w_i H_0(i) = 67.57 \pm 0.52\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}. \tag{7}$$

Hence the agreement of $\langle H_0 \rangle$ with (6), apart from the uncertainties, is again excellent (0.3%). (Note: The uncertainties for $H_0 = 67.78^{+0.91}_{-0.87}$ have been rounded to $H_0 = 67.78 \pm 0.90$). As remarked earlier, the CMB and BAO determinations are based on distance determinations that do not explicitly involve the deceleration parameter, unlike the LDL determinations that do, and hence they can be alternatively attributed to a decelerating model that gets the same distance as the ΛCDM model to within experimental error. However, this simple observation seems to fail, in view of the recent HOLiCOW XIII result of Wong *et al.* [25], that finds for a flat ΛCDM model, the value, $H_0 = 73.3^{+1.7}_{-1.8}\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$, based on a joint analysis of six gravitationally lensed quasars with measured time delays. Since the different time delays result from different distances traversed by light rays following different paths from the source to the observer, one is dealing with what is essentially a distance measurement, and therefore, from the above argument, the results should have

agreed with the CMB and BAO values, instead of agreeing with the LDL value, as they do. This apparent contradiction with the proposed model will be taken up in Section 5, where the implications of the various models for the age of the universe are discussed.

Finally, it should be mentioned that from the gravitational wave (GW) standard siren measurement of the distance of the binary neutron star merger [26] [27] that was the source of GW170817, combined with optical identification of the host galaxy NGC 4993 that provided the redshift determination, it was possible to determine a value of the Hubble constant given by $H_0 = 70^{+12}_{-8}\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$, as described in Abbot *et al.* [28]. In determining H_0, they used $v_H = H_0 D$, where v_H is the Hubble flow velocity, D is the distance to the source, and $v_H = cz$. Since for this case $z \approx 0.01$, they could ignore higher order corrections involving $k(z, q_0, j_0)$, that would have contributed of the order of one percent. More recently, a less uncertain determination of H_0 has been made by Hotokezaka *et al.* [29] by using the high angular resolution imaging of the radio counterpart of GW170817, and combining this with previous GW and electromagnetic (EM) data. They found that $H_0 = 70.3^{+5.3}_{-5.0}\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$. These two results for H_0 are too uncertain to decide between the CMB and BAO lower values, and the LDL higher values of H_0, although, as H.-Y. Chan *et al.* [30] pointed out, future GW detection by LIGO and VIRGO can be expected to improve the determination of H_0 to a precision of approximately two percent in five years, and approximately one percent within ten years, and this should enable a judgment between the two discordant values to be made. According to the decelerating model, since one is dealing with a distance measurement, the result should agree with the CMB and BAO values, or the model would be wrong. However, it should be noted that the low velocity relation $v_H = H_0 D$, that they used to determine H_0 depended on the fact that the redshift of the source galaxy was so small that the higher order corrections involving $k(z, q_0, j_0)$ could be neglected. But if some of the future GWs are from binary neutron star mergers that are at higher redshifts, say, $z \approx 0.07$, so that it would be necessary to include the next order correction, $k(z, q_0) = \left(1 + \frac{1}{2}(1 - q_0)z\right)$, then the determination of H_0 would have to include q_0, and if they set, $q_0 = -0.53$, in accordance with the current accelerating universe paradigm, the result should agree with the value of H_0 found in [13], allowing for the uncertainties in both determinations. But if instead it should agree with the SHoES value, then it would also be another indication that the decelerating model proposed here is wrong. Thus, standard siren measurements can play a fundamental role, not only in determining H_0, but in determining whether a major paradigm shift from the current accelerating universe to a decelerating universe is needed. This view is in contrast with alternative explanations for the Hubble constant disagreement, such as in [31] [32] [33] [34], and many others, that rely on *ad hoc* corrections to the current accelerating model, and hence do not predict that the present disagreement about the Hubble constant requires a major paradigm shift to a decelerating universe.

3. Consequences of Galactic Clusters for the Model

In the model that has been developed up until now galactic clusters were neglected for simplicity. An additional purpose of this work is to revise the model by taking this clustering into account, and to examine some of the resulting consequences. One such consequence will be that it explains why the difference of arrival times between neutrinos and photons from SN1987a [35] [36] [37] was only three hours, see ref.1 in [35], rather than thousands of years, as would have been the case if all intergalactic space (such as that between the Milky Way (MW) and the Large Magellanic Cloud (LMC) in which SN1987a was located) necessarily contained dark energy with an index of refraction of $n \approx 1.5$, for $z \le 0.6$. Indeed, a main consequence of this revised model is that it incorporates the fact that some galactic clusters have so much dark matter holding them together that they do not partake of the Hubble expansion. In this regard, these clusters behave like normal galaxies. To be sure, this recognition traces back to the pioneering work of Zwicky [38] with his well-known studies of galactic peculiar velocities in the Coma cluster, and his recognition that the cluster contained more gravitating mass than one could infer from the observed light, as well as to subsequent numerous large-scale studies, well beyond the scope of this work. Galactic clusters that have such a strong gravitational binding because of their dark matter content that they do not experience the Hubble expansion will be called here "tight clusters." In contrast, superclusters of galaxies, that are so spread out that they experience the Hubble expansion, might be referred to as "loose clusters." It follows that, in accordance with the original model, the intergalactic space within tight clusters does not have any dark energy, for which $n > 1$, since, according to the model, it is the Hubble expansion after the epoch corresponding to $z = 1.65 \pm 0.15$ [15] that caused the dark matter that was in the expanding space to undergo a phase transition into dark energy, with its associated increase of the index of refraction. A relevant case of a tight cluster for the considerations here is the Local Group that contains the MW. That the Local Group is a tight cluster follows, for example, from the well-known fact that the great spiral galaxy in Andromeda, M31, at a distance of ~2.5 Mlyr, is approaching the MW rather than receding, as would be the case if it were experiencing the Hubble expansion. It follows that the LMC, at ~1.5×10^5 lyr from the MW, since it is well within the Local Group, is not within a sea of dark energy, but only within the tight cluster's dark matter, for which $n = 1$. Therefore, the photons that came to the earth from the explosion associated with SN1987a traveled with speed c, and hence the enormous time delay did not occur, that would have occurred if they had traveled through dark energy from the LMC to the MW at a speed of approximately $2c/3$. Although this was the explanation for the absence of a lengthy time delay that was given previously in [19], it was *ad hoc*, since it was not justified, as is done here, on the general grounds that the LMC and MW are located in a tight cluster, the Local Group. If this model is correct, tight clusters must have a boundary region where n varies from unity within the cluster to $n > 1$ outside the cluster, depending on the redshift of the cluster. The shape

and depth of this boundary where dark matter is undergoing a phase transition into dark energy, with the associated increase in the index of refraction, is of course unknown. Superclusters would not have such a boundary region. For $z >\sim 1.7$, since the phase transition of the dark matter into dark energy has not yet occurred, it follows that at these higher redshifts the speed of light within superclusters is the vacuum speed of light c, as it is in tight clusters. Consequently, the speed of light is then c throughout all space, since, in accordance with the model, there is no dark energy at all, and the speed of light through the dark matter that would then be filling space in its place is c.

As was first noted in [8], and further discussed in [39] (which contains numerous references to the literature), as is well-known, because of their very low masses, for the energies under consideration, neutrinos travel exceedingly close to c, while according to the model, gamma ray bursts (GRBs), when traveling over distances with redshifts $z < 1.7$, but outside of tight clusters through which they might pass on their journey to the MW, travel at speeds less than c, and hence there should not be any coincidences between the arrival of neutrinos and GRBs. Until fairly recently, no such coincidences had been observed, as referenced in [39], when, as was noted in [12], on 22 September 2017, as reported in Science (18 July, 2018), by Aartsen *et al.* [14], there was observed at IceCube a neutrino 170922A at 290 TeV that came to within 0.1 degrees from the direction of the blazar TXS-0506+056, at a redshift $z = 0.34$. The blazar was reported as flaring at the time with enhanced activity in the GeV range. The correlation between the neutrino and the flaring blazar was given as 3σ. If there had been a correlation of 5σ, the model proposed here would have been proven wrong. Thus, there is still the possibility that the model can survive this observation, since to date, no new examples of possible coincidences of neutrinos with GRBs have been observed. However, interestingly, a study of earlier possible coincidences with the same blazar has been found by the IceCube Collaboration [40]. They reported that two years prior to the above discovery, by examining 9.5 years of archival data, there was an independent 3.5σ excess of neutrino flux from the direction of the same blazar. Once again, if these neutrinos did indeed actually come from this blazar, the proposed model would be wrong. However, there is a difficulty with drawing such a conclusion: Active galactic nuclei are widely distributed across the sky [41], so that if blazars are the sources of coincidental GRBS and neutrinos, it is difficult to understand why the observation of possible coincidences has only shown up for the blazar TXS-0506+056. On the other hand, if the observed neutrinos, both from the 2017 detection, and from the archival data, were actually from a more distant source that by chance happened to be located, angle-wise, near to blazar TXS-0506+056, so that there were accidental correlations, this would explain why one has not seen possible coincidences of neutrinos and GRBSs associated with other blazars. Thus, until such possible coincidences with other blazars are observed, it is reasonable to conclude that the proposed model is still viable. Finally, it should be noted that ga-

lactic clusters impact on the previous discussion of a supplementary source of discordant redshifts given in [10] [39]. Thus, if the main discordant redshift galaxy is in a tight cluster, the refractive explanation given previously no longer holds. However, if the main discordant redshift galaxy is outside of a tight cluster, or within a supercluster, so that it is surrounded by dark energy, the previous discussion is still valid, although, since there is now in the revised model a variable index of refraction, as will be discussed in the next section, this supplemental source of discordant redshifts should diminish as $z \to 1.7$, since the index of refraction approaches unity at that value.

4. A Variable Index of Refraction

In the previous work it was assumed that for the redshift range $0 \le z \le 0.6$, the index of refraction of the dark energy n is a constant, and it was found that a least squares fit to the ΛCDM model for various values of Ω_m yielded the following range of values for n given by $1.47 \le n \le 1.54$. For the values $1.47 \le n \le 1.50$, the values of the distances predicted by the model for the redshift range $0.1 \le z \le 0.5$ were less than that for the ΛCDM model, whereas, as discussed in Section 2, the model requires that the distances should be greater than that for the ΛCDM model, when the latter is combined with the LDL determinations of H_0. On the other hand, for $n = 1.54$, as shown in Table 2 in [9] at $z = 0.1$, $z = 0.2$ (note the range of redshifts for LDL in [11] was $0.023 < z < 0.15$), the model predicts a greater distance than the ΛCDM model, so that the choice of a constant value of $n = 1.54$ would seem to be favored. However, this is ruled out, since it predicts progressively even greater distances than ΛCDM at higher redshifts, whereas ΛCDM fits very well at these higher redshifts, particularly at $z = 0.5$, the value of redshift where the majority of SNe Ia were located that led to the accelerating ΛCDM model [4] [5] [6]. In fact it was found that for $\Omega_m = 0.315$, the value of the constant index of refraction model needed to fit ΛCDM at $z = 0.5$ is $n = 1.46$. Thus it is clear that to satisfy the above requirements, a constant index of refraction will not suffice. This is further emphasized by the fact that n must go to unity at $z = 1.7$, where Riess *et al.* [15] found no evidence of acceleration, as noted above, and this is interpreted in the model as being where the Hubble expansion caused the dark matter outside the tight clusters and cluster-free galaxies to start to undergo a phase transition into dark energy. Note that, for simplicity, the value of $z = 1.65 \pm 0.15$ in [15] has been rounded to $z = 1.7$. Also, since it was found in [4] [5] [6] that there was no dispersion over the range of wavelengths that were observed, it is assumed, as was done earlier, that the index of refraction is without dispersion over all wavelengths, as it is for the vacuum, both in special relativity and general relativity. In this respect, the index of refraction in this model differs significantly from its behavior in standard electromagnetic theory.

The above considerations lead to the conclusion that the revised model should introduce a variable index of refraction $n(z)$ for $0 \le z \le 1.7$, and a constant

index $n(z)=1$, for $z \geq 1.7$. The behavior of the function $n(z)$ is determined by the conditions: $n(1.7)=1$, and $n(0.5)$ should have the value required to fit the ΛCDM model at the redshift $z=0.5$, for the present density values $\Omega_m = 0.315$, $\Omega_\Lambda = 0.685$. To arrive at a possible expression for $n(z)$, it is convenient to review the original argument in [7] that led to the expression for the increase in apparent magnitude given by $\delta m = 5\log(1+(n-1)\ln(1+z))$, and the related increase in logarithm of distance given by $d = \log(1+(n-1)\ln(1+z))$, for a constant index of refraction. There, one derived the increase of the apparent magnitude, due to the reduction in the speed of light, by first obtaining the infinitesimal extra amount of time δt it takes light to travel an infinitesmal distance $\mathrm{d}\sigma$ through the dark energy with a dispersionless index of refraction *n*, compared with that through the vacuum with its dispersionless index of refraction of unity. One has that through the dark energy, the infinitesimal time $\mathrm{d}t'$ for light to travel an infinitesimal distance $\mathrm{d}\sigma$ is $\mathrm{d}t' = n\mathrm{d}\sigma/c$, while that through the vacuum is $\mathrm{d}t = \mathrm{d}\sigma/c$, hence $\delta t = \mathrm{d}t' - \mathrm{d}t = (n-1)\mathrm{d}\sigma/c = (n-1)\mathrm{d}t$. During that infinitesimal extra amount of time the universe will have expanded by the infinitesimal extra amount $\dot{a}\delta t = (n-1)\dot{a}\mathrm{d}t$, and therefore the infinitesimal fractional amount of extra expansion is

$$(n-1)\dot{a}\mathrm{d}t/a = (n-1)\mathrm{d}\ln a. \tag{8}$$

This expression is to be integrated from the time of emission of the light to the present, in order to obtain the total fractional amount of extra expansion. However since it is redshift rather than time that is measured, it is appropriate to convert this relation to one that is a function or redshift using $a = a_0/(1+z)$, which holds for arbitrary FLRW expanding universes. Under this assumption, and the assumption that *n* is now a function of redshift, $n = n(z)$, the integral of (8) takes the form

$$\int_0^z (n(z')-1)\mathrm{d}\ln(1+z'). \tag{9}$$

Now in [7], since *n* was assumed to be a constant, the factor $(n-1)$ was removed from under the integral sign to obtain the value of the integral as $(n-1)\ln(1+z)$, but here it will be assumed, for simplicity, that

$$n(z) = n_0 - bz, \quad 0 \leq z \leq 1.7, \quad n(z) = 1, \quad z \geq 1.7. \tag{10}$$

Under this assumption, (9) takes the form

$$\int_0^z (n_0 - bz' - 1)\mathrm{d}\ln(1+z'). \tag{11}$$

The value of this integral is

$$(n_0 - 1 + b)\ln(1+z) - bz, \tag{12}$$

so that the original luminosity distance D_L is increased by the factor $(1+(n_0-1+b)\ln(1+z)-bz)$, and hence the resulting luminosity distance D_L' is given by

$$D_L' = (1+(n_0-1+b)\ln(1+z)-bz)D_L. \tag{13}$$

Then, following the argument on page 82 of [7], for the apparent magnitude δm, and for the increase in the logarithm of distance d, one has

$$\delta m = 5\log\left(1+\left(n_0-1+b\right)\ln\left(1+z\right)-bz\right), \tag{14}$$

$$d = \log\left(1+\left(n_0-1+b\right)\ln\left(1+z\right)-bz\right). \tag{15}$$

Note that in the limit $b\to 0$, the above expressions reduce to the previous ones for a constant index of refraction, upon setting $n_0 = n$, and note also that $\delta m = 5d$. The parameters n_0 and b are next determined from the requirements described above. Thus from (10), one has

$$n_0 - b1.7 = 1, \tag{16}$$

and at $z = 0.5$, one has

$$\log\left(\mathrm{X}_\Lambda\left(0.5\right)/\mathrm{X}_m\left(0.5\right)\right) = d\left(0.5\right) = \log\left(1+\left(n_0-1+b\right)\ln\left(1.5\right)-b\left(0.5\right)\right), \tag{17}$$

where $\mathrm{X}_\Lambda, \mathrm{X}_m$ are defined and worked out in Equations (17), (18) in [7], except that there it was for the case that $\Omega_m = 0.3, \Omega_\Lambda = 0.7$, so that $\Omega_\Lambda/\Omega_m = 2.3333$, whereas here, using the more recent rounded values $\Omega_m = 0.315, \Omega_\Lambda = 0.685$, and hence $\Omega_\Lambda/\Omega_m = 2.1746$. For these new values, for Equations (17), (18) in [7], one finds the following replacements

$$\mathrm{X}_\Lambda = \left(0.315\right)^{-1/2}\int_0^{0.5}\left(\left(1+z\right)^3+2.1746\right)^{-1/2}\mathrm{d}z = 0.4387, \tag{18}$$

where the integral has been evaluated numerically, while by direct integration one has

$$\mathrm{X}_m = \int_0^{0.5}\left(1+z\right)^{-3/2}\mathrm{d}z = 0.3670, \tag{19}$$

hence $\log\left(\mathrm{X}_\Lambda\left(0.5\right)/\mathrm{X}_m\left(0.5\right)\right) = 0.0775 = d\left(0.5\right)$. Since from (16) $n_0 - 1 = b1.7$, upon substitution in (17), the following relation for b results

$$\log\left(1+2.7b\ln\left(1.5\right)-0.5b\right) = 0.0775, \tag{20}$$

from which one obtains, with the aid of (16), the following values for the two parameters

$$b = 0.3285,\quad n_0 = 1.558. \tag{21}$$

Obviously, the above values are only approximate because of the uncertainties in Ω_m and the value of redshift for which $n = 1$. It will be noted that for $z = 0.5$, from (10) and (21) one has that $n\left(0.5\right) = 1.394$. This is significantly lower than the values for constant n that was found earlier, depending on the values of Ω_m, to be in the range 1.47 - 1.5. It is interesting to compare the values of $\log\left(\mathrm{X}_\Lambda/\mathrm{X}_m\right)$ with the values of d, as was done in earlier papers, to see how well this model fits the ΛCDM model. This is done in **Table 1** that covers a greater range of redshifts than was done in these earlier papers, so as to include much of the range in [13], and also to include some beyond $z = 1$, where future astronomical observations of SNe Ia will provide more data, and opportunity for comparison. It will be noted that in **Table 1** $d(1.7)$, $d(1.9)$, and $d(2.1)$ are all equal, this is because for $z \geq 1.7$, $n\left(z\right) = 1$, and hence the contribution to the

Table 1. In columns 2 - 5, comparison of $\log(X_\Lambda/X_m)$ with $d \equiv \log\left(1+(n_0-1+b)\ln(1+z)-bz\right)$ for $n_0 = 1.558$, $b = 0.3285$, $\Omega_m = 0.315$, $\Omega_\Lambda/\Omega_m = 2.1746$, $\Delta \equiv d - \log(X_\Lambda/X_m)$, and $R_\Lambda \equiv \log(X_\Lambda/X_m)$.

z	$\log(X_\Lambda/X_m)$	d	Δ	Δ/R_Λ %
0.005	0.00110	0.00121	0.00011	10.0
0.02	0.00433	0.00470	0.00037	8.5
0.04	0.00864	0.00930	0.00066	7.7
0.06	0.01273	0.01367	0.00094	7.4
0.08	0.01675	0.01786	0.00111	6.6
0.10	0.02061	0.02189	0.00128	6.2
0.30	0.05330	0.05467	0.00137	2.6
0.50	0.07752	0.07752	0	0
0.70	0.09580	0.09367	−0.00213	-2.2
0.90	0.10989	0.10505	−0.00484	-4.4
1.10	0.12095	0.11284	−0.00811	-6.7
1.30	0.12979	0.11783	−0.01196	-9.2
1.50	0.13698	0.12056	−0.01642	-12.0
1.70	0.14292	0.12140	−0.02152	-15.1
1.90	0.14790	0.12140	−0.02650	-17.9
2.10	0.15212	0.12140	−0.03072	-20.2

integral (11) from $z = 1.7$ to $z = 2.1$ vanishes.

In concluding this section, it should be emphasized that the above linear model for the index of refraction is only a preliminary attempt to go beyond the constant index of refraction of the previous publications. It is possible a deeper understanding of dark matter and its proposed phase transition to dark energy will emerge from the comparison of the above predictions with that of the ΛCDM model, as well as with other accelerating models, that will lead to an improved model for the behavior of the index of refraction.

5. Age of the Universe

The age of the universe, *i.e.* the time back to the Big Bang where the expansion parameter $a(t)$ was arbitrarily small, was first taken up in [8], where it was shown that although the EdS universe leads to an age T_0 given by $T_0 = (2/3)H_0^{-1}$, which would rule out the model, the reduction in the speed of light, occasioned by the dark energy that has a constant index of refraction n, leads to an age $T_0 = (2/3)nH_0^{-1}$, and since in the original model $n \approx 3/2$, this well-known objection to the EdS universe was removed. However, before examining what the revised model with a variable index of refraction has to say about this issue, it will be helpful to recapitulate and comment on how the current disagreement about

the Hubble constant impacts on the age of the universe, as well as the latter, reciprocally, on the Hubble constant.

The recent Planck CMB rounded value of the Hubble constant is $H_0 = 67.4 \pm 0.5\ \mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$ [20]. Likewise, Planck gives the age of the universe as 13.797 ± 0.023 Gyr, which will be replaced by $T_0 = H_0^{-1} F$, where $F = 0.95$, as will be derived below, and since $H_0^{-1} = 14.5 \pm 0.1\ \mathrm{Gyr}$, the rounded age for Planck CMB is $T_0 = 13.8 \pm 0.1\ \mathrm{Gyr}$. There will be further comment on the meaning of this result below. To derive the above value of *F*, one makes use of Einstein's field equation for the energy density

$$G_0^0 = -\kappa T_0^0 - \Lambda, \tag{22}$$

where $\kappa \equiv 8\pi G/c^4$. For $T_0^0 = \rho c^2$, where ρ is the matter density consisting of the dark matter mass density and the baryonic mass density, while $\kappa^{-1}\Lambda$ is the density of dark energy in the ΛCDM model. This is the simplified two component energy density source model for ΛCDM that has been used throughout this work. Upon introducing the FLRW line element for a flat universe (analysis for a slightly closed universe is in [9]), one has

$$\mathrm{d}s^2 = c^2 \mathrm{d}t^2 - a(t)^2 \delta_{ij} \mathrm{d}x^i \mathrm{d}x^j, \ i, j = 1, 2, 3. \tag{23}$$

After the metric is substituted in G_0^0, and suitably rewritten,(22) takes the form

$$\frac{\dot{a}^2}{2} - \frac{4\pi G \rho a^2}{3} - \frac{\Lambda a^2 c^2}{6} = 0. \tag{24}$$

Since the matter source tensor T_ν^μ is that for cold dark matter $T_\nu^\mu = diag\left(T_0^0, 0, 0, 0\right)$, there is no pressure. It follows from the covariant conservation law $T_{\nu\ ;\mu}^{\mu} = 0$ that ρa^3 is a constant of the motion. Hence, upon introducing the mass $M \equiv 4\pi \rho a^3/3$, (24) may also be written as

$$\frac{\dot{a}^2}{2} - \frac{GM}{a} - \frac{\Lambda a^2 c^2}{6} = 0. \tag{25}$$

In this form, one recognizes (25) as the Newtonian equation for a test particle moving radially outside a fixed spherical body of mass *M* with kinetic energy per unit mass $T = \dot{a}^2/2$, and potential energy per unit mass $V = -(GM/a) - \Lambda a^2 c^2/6$, so that the total energy per unit mass is $T + V = 0$. The fact that the mass of the test body cancels out completely from the equation can be seen as a manifestation of the Newtonian principle of equivalence, inertial mass equals gravitational mass, which is not only obeyed by the standard gravitational potential energy term, in accordance with Newtonian mechanics, but by the cosmological potential energy per unit mass term as well. One sometimes refers to the cancelled mass that would multiply *M* as the "passive" gravitational mass, while *M*, since it comes from the source tensor T_ν^μ, is described as the "active" gravitational mass. However, since the source energy-momentum tensor exists in special relativity where there is no gravitation, *M* is necessarily inertial mass. But since it also has been shown that inertial mass is passive gravitational mass, it follows

that all three masses are the same in general relativity, as they are in Newtonian mechanics. Thus, at this level, general relativity has a deep connection with Newtonian mechanics, so that if general relativity were to break down, as has been occasionally proposed as an alternative explanation for the diminished brightness of the SNe Ia, Newtonian mechanics would also have to break down as well, at its most basic level, something that is frequently not discussed in such proposals.

Returning now to the derivation of F which, although standard, is given here for completeness. From (25), after solving for $\dot{a}$, and taking the positive root, one has

$$\dot{a}/a = \left(\left(8\pi G\rho/3\right)+\left(\Lambda c^2/3\right)\right)^{1/2}. \tag{26}$$

The negative root, not indicated above, corresponds to the descending branch of the ΛCDM universe, since, as discussed in conjunction with the EdS universe in [9], to be consistent with the time-reversal invariance of the field equations, ΛCDM should be thought of as a cyclic universe with infinite period, and infinite amplitude, with a cusp at $t=0$ where the collapsing part of the cycle meets the expanding part; although this feature of the ΛCDM universe is rarely discussed. However, for the purpose of this work, one is only concerned with the positive root, corresponding to the expanding portion of the cycle. Upon introducing the parameter $H \equiv \dot{a}/a$, the Hubble constant $H_0 \equiv \dot{a}_0/a_0$, and using the relation $a = a_0/(1+z)$, so that $\rho = \rho_0(1+z)^3$, and with the following definitions, $\rho_c \equiv 3H_0^2/8\pi G$, $\rho_\Lambda \equiv \Lambda c^2/8\pi G$, $\Omega_m = \rho/\rho_c$, and $\Omega_\Lambda \equiv \rho_\Lambda/\rho_c$, (26) takes the form

$$H/H_0 = \Omega_m^{1/2}\left(\left(1+z\right)^3+\left(\Omega_\Lambda/\Omega_m\right)\right)^{1/2}. \tag{27}$$

Since $\mathrm{d}t = H^{-1}\mathrm{d}a/a$, it follows, after integrating, that the age of the universe is given by

$$T_0 \equiv \int_0^{T_0}\mathrm{d}t = H_0^{-1}\Omega_m^{-1/2}\int_0^{a_0}\left(\left(1+z\right)^3+\left(\Omega_\Lambda/\Omega_m\right)\right)^{-1/2}a^{-1}\mathrm{d}a. \tag{28}$$

Then, upon using $a^{-1}\mathrm{d}a = -a_0\mathrm{d}z/(1+z)$, and $T_0 = H_0^{-1}F$, one has

$$F = \Omega_m^{-1/2}\int_0^{\infty}\left(\left(1+z\right)^3+\left(\Omega_\Lambda/\Omega_m\right)\right)^{-1/2}\left(1+z\right)^{-1}\mathrm{d}z. \tag{29}$$

Since $\Omega_\Lambda/\Omega_m = 2.1746$, as noted preceding (18), and $\Omega_m^{-1/2} = 1.7817$, (29) yields $F = 0.951$. Since, more accurately, one has $\Omega_m = 0.315\pm0.017$ [20], (29) yields $F\left(\Omega_m = 0.308\right) = 0.957$, and $F\left(\Omega_m = 0.322\right) = 0.945$, so that

$$F = 0.951\pm0.006. \tag{30}$$

However, in view of the uncertainty in H_0, and the simplification of a two component flat ΛCDM model, it is clearly appropriate to set $F = 0.95$ in what follows.

If one now determines the age of the universe for the SHoES LDL value of H_0 given by Riess *et al.* [20] as $H_0 = 74.03\pm1.42\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$, that will be rounded here to $H_0 = 74.0\pm1.4\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$, so that $H_0^{-1} = 13.2\pm0.2\ \mathrm{Gyr}$,

one has for the age of the flat ΛCDM universe

$$T_0 = H_0^{-1}F = 12.6 \pm 0.2 \text{ Gyr}. \tag{31}$$

This age is in possible conflict with the ages of the globular clusters in the MW, for which the cut-off as to lower age was given as 12.6 Gyr by Krauss and Chaboyer [42]. While quite recently, Gellart *et al.* [43], using color-magnitude diagrams, and the new distances to stars in the MW found by Gaia [44], obtained a peak age for stars of ~13.4 Gyr. Here the disagreement of the stellar ages with the SHoES LDL age in (31) is even greater than it is for the globular clusters. Nevertheless, this short age in (31) would seem to have found support by the above mentioned findings of HOLiCOW XIII [25] that, in contrast with the LDL method based on Cepheids, used a time delay system based on gravitationally lensed quasars, as given in Wong *et al.* [25]. Since HOLiCOW XIII found $H_0 = 73.3^{+1.7}_{-1.8}\ \text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$, this results in a Hubble time $H_0^{-1} = 13.3 \pm 0.4$ Gyr, and under the assumption that this value, as with the SHoES LDL value, is for a flat ΛCDM universe, so that with $F = 0.95$, one has for HOLiCOW's age of the universe

$$T_0 = 12.7 \pm 0.4 \text{ Gyr}. \tag{32}$$

It is in good agreement with (31), and again in possible disagreement with the above stellar ages. However, there is another issue arising with HOLiCOW's determination of H_0. Since their determination involves a *distance determination of* H_0, and according to the arguments presented earlier in Section 2, it should agree with the CMB and BAO determinations to within the uncertainties, which it obviously does not, and hence, *either the decelerating model proposed here is wrong, or there is possibly some problem with HOLiCOW's determination of* H_0. To examine this, it will be noticed that their value for H_0 is based on six independent determinations that can be divided into two groups, A and B. Group A has four members (units, $\text{kms}^{-1}\cdot\text{Mpc}^{-1}$, in what follows are omitted for brevity): $68.9^{+5.4}_{-5.1}, 71.0^{+2.9}_{-3.3}, 71.6^{+3.8}_{-3.9}, 71.7^{+4.8}_{-4.5}$, which have been simplified to those given in column A below, since more accuracy is not needed for the analysis. The second group B has two members: $78.2^{+3.4}_{-3.4}, 81.1^{+8.0}_{-7.1}$, which again have been simplified in column B, and so one has

A	B
69.8 ± 5.3	78.2 ± 3.4
71.0 ± 3.1	81.1 ± 7.6
71.6 ± 3.9	
71.7 ± 4.7	

(33)

Before proceeding further, it is desirable to re-evaluate H_0 using the data in the above columns, and the weighted average method used earlier, to see how it compares with their more precise determination of H_0. It was found that $\langle H_0 \rangle = 73.3 \pm 2.0$, which is in perfect agreement with the value of H_0 obtained by HOLiCOW, given by $73.3^{+1.7}_{-1.8}$ and is therefore in more than sufficient agreement with their value to justify the analysis that follows. It was found that

the weighted average of the values in column A yielded $\langle H_0(\mathrm{A})\rangle = 71.1 \pm 2.2$, while that in column B yielded $\langle H_0(\mathrm{B})\rangle = 78.7 \pm 4.2$. One sees that the difference, ignoring the uncertainties in the mean values, $\langle H_0(\mathrm{B})\rangle - \langle H_0(\mathrm{A})\rangle = 7.6$ is significantly greater than the difference $\langle H_0(\mathrm{A})\rangle - H_0(\mathrm{CMB}) = 3.7$, so that one may question combining the data in B with that in A. A further objection is the age of the universe implied by $\langle H_0(\mathrm{B})\rangle$, since one has $(\langle H_0(\mathrm{B})\rangle)^{-1} = 12.5 \pm 0.7$ Gyr, and, if this is to be associated with the flat ΛCDM universe, then upon multiplication by F = 0.95, one has $T_0(\mathrm{B}) = 11.9 \pm 0.7$ Gyr, an age range that is hardly acceptable for the globular clusters, let alone the oldest stars in the MW. This suggests that in determining the value of the Hubble constant by HOLiCOW III, one should reject the data from column B, and base the determination solely on that from column A, and clearly the small value for the difference $\langle H_0(\mathrm{A})\rangle - H_0(\mathrm{CMB})$ at 1.6σ, does not disprove the proposed decelerating model.

It is now appropriate to examine the age of the decelerating universe predicted by the revised model, in which the index of refraction is no longer a constant n, but a variable $n(z)$. It will be assumed that the age takes the form

$$T_0 = (2/3)\langle n\rangle H_0^{-1}, \tag{34}$$

where $\langle n\rangle$ is the average of the index of refraction over a suitable redshift distance ζ that is determined from the integral

$$\langle n\rangle = \zeta^{-1}\int_0^{\zeta} n(z)\mathrm{d}z. \tag{35}$$

To obtain <n>, it will be assumed (to be discussed further below) that just as the CMB determines a Hubble constant for a decelerating universe, the age that it predicts is also for the decelerating universe, hence

$$T_0 = (2/3)\langle n\rangle H_0^{-1} = 0.95H_0^{-1}, \tag{36}$$

so that $\langle n\rangle = 1.43$. In evaluating the above integral, it is convenient to break it up into two parts: The first part extends from the terrestrial observer to the edge of the local group, since $n(z) = 1$ over this range. For simplicity, this redshift distance will be taken to be the redshift ℓ determined by the diameter of the local group, since it is going to prove to be negligible. To be sure, since the local group is a tight cluster, it is not expanding, but one can imagine a galaxy, with negligible peculiar velocity, located just outside the local group, that is experiencing the Hubble expansion at the redshift ℓ. To obtain a value for ℓ it is convenient to use the expression for the first order Doppler effect $cz/\langle n\rangle = H_0 D_L$, since the desired value of redshift is so small, higher order terms may be neglected. Now, with D_L taken to be the diameter of the local group, that is estimated to be ~10^7 lyr [45], and with the value, $H_0 = 67.4\ \mathrm{km\cdot s^{-1}\cdot Mpc^{-1}}$ from (20), one obtains an upper bound for ℓ given by $\ell = 10^{-4}$. Upon performing the integration in (34) with $n(z) = 1, 0 \le z \le \ell$, and with $n(z) = n_0 - bz, \ell \le z \le \zeta$, (34) becomes $\langle n\rangle = \ell\zeta^{-1} + n_0 - n_0\ell\zeta^{-1} - (b/2)\zeta + (b/2)\ell^2\zeta^{-1}$. Upon multiplication of both sides equation by ζ, and rewriting it, the following quadratic equ-

ation for ζ results

$$\zeta^2 - 2b^{-1}\left(n_0 - \langle n\rangle\right)\zeta + 2b^{-1}\ell\left(n_0 - 1 + (b/2)\ell\right) = 0. \tag{37}$$

After inserting the numerical values, $n_0 = 1.56$, $\langle n\rangle = 1.43$, $b = 0.3285$, to sufficient accuracy, (36) becomes

$$\zeta^2 - 0.79\zeta + 3.75\times 10^{-4} = 0. \tag{38}$$

The larger root $\zeta = 0.79$ is clearly the root of physical interest. It is significant that ζ exceeds the redshift z = 0.5, where the model and ΛCDM are in perfect agreement (by selection) with the diminished brightness of the SNe Ia. On the other hand, as noted in the previous section, the function $n(z)$ should emerge from a theory that goes beyond the present model. It is possible in such a development ζ will also emerge, and hence one will be able to use (34) to obtain $\langle n\rangle$, and compare it with the value $\langle n\rangle = 1.43$ found above. Also, a matter for further investigation is the assumption that the age of the decelerating universe is $0.95H_0^{-1}$, where H_0 is the CMB Hubble constant that is claimed here to be that for a decelerating universe, since, as was shown above, F = 0.95 was obtained for the flat ΛCDM accelerating universe. Thus, rigorously, the age 13.8 ± 0.1 Gyr is actually the age of an accelerating universe that presumably has the same age as that of the decelerating universe proposed here. However, under the assumption the distances used in obtaining the Hubble constant found in [13] are correct, in contrast with the SHoES value, the age of the currently proposed ΛCDM accelerating universe is

$$0.95H_0^{-1} = 0.95\left(69.8 \pm 2.0 \text{ km/s/Mpc}\right)^{-1} = 13.2 \pm 0.2 \text{ Gyr}. \tag{39}$$

As was shown in previous discussions, the age of the decelerating universe is greater than the age of the accelerating universe, consequently, while the age 13.8 ± 0.1 Gyr used here for the decelerating universe is *qualitatively* correct, since it is greater than the above age of the accelerating universe, it will require further theoretical study to determine whether the use of F = 0.95 in (36) is justified in obtaining the age of the decelerating universe.

Meanwhile, just as this work was nearing completion, it was reported that the Atacama Cosmology Telescope (ACT) group [46], by studying the polarization in the CMB, were led to values of the Hubble constant and the age of the universe sufficiently close to those found by the Planck collaboration that no changes in the above considerations are needed.

6. Conclusions

Undoubtedly, the major support for the model, so far, is the finding in Section 2 that when $q_0 = -0.53$ for an accelerating universe, that was used by Freedman *et al.* [13] in obtaining their value of H_0, is replaced by $q_0 = 0.50$ for a decelerating universe, and the distance they determined is assumed to be the same as that required for the model for the redshift $z = 0.07$, then their value of H_0 comes down to excellent agreement (0.1%) with the CMB value, and likewise to

within 0.3% of the weighted average of the BAO values for H_0, despite the uncertainties of a few to several percent in these different measured values of H_0. This result partially validates the claim underlying this work that the disagreement about the Hubble constant stems mainly from the proposal that the CMB and BAO values for H_0 are actually for a decelerating universe, rather than for an accelerating universe, as is currently believed. However, to fully validate the claim, it remains to be seen what value Freedman *et al.* [13] obtain for H_0 when its re-evaluation has been made for the full range of redshifts that were used in obtaining their present result.

Additional support for the model stems from the fact, as discussed in Section 3, that there are tight clusters, such as the local group, of which the Milky Way and the LMC are members, in which the gravitational binding of the dark matter content is so strong that these clusters do not undergo the Hubble expansion, and, as a consequence, in accordance with the model, their dark matter does not undergo a phase change into dark energy, so that the speed of light within these tight clusters remains c, as it does within galaxies themselves. This leads to a simple explanation as to why the light from SN1987a in the LMC was able to arrive 3 hours after the neutrinos that were detected. On the other hand, the reduction in the speed of light that is traveling through the dark energy of the space that is experiencing the Hubble expansion leads to the prediction that there should not be any correlation between neutrinos and gamma ray bursts (GRBs). As reported in [1] and more fully in [39], this turned out to be well satisfied in thousands of cases for low MeV neutrinos, and about a hundred cases of TeV neutrinos, and a few cases for PeV neutrinos. However, there is a possible exception in the TeV range that was not referenced in [1]. This is the case of a neutrino that at the 3σ level seems to have been emitted from the blazar TXS-0506+56 that was flaring at the time. Also a study of past records indicates there appears to have been other possible neutrino emissions from the same blazar. It is suggested in the text that these cases are possibly accidental correlations associated with a more distant neutrino source that happens to lie within 0.1 degrees, angle-wise, near the blazar. But if this should prove not to be the case, this would be a strong indication that the model is wrong, despite the support described in Section 2.

In Section 4, an attempt was made to improve on the original assumption of a constant index of refraction by introducing a variable index of refraction $n(z)$ that was taken to be linear for simplicity. However, this complicated the determination of the age of the universe, as discussed in Section 5, and clearly more study of this issue is needed. On the other hand, the analysis showed that the higher values of the Hubble constant found by different groups is problematic, since their values lead to ages of the universe in possible conflict with the ages of the oldest stars in the Milky Way.

In conclusion, as noted earlier in [39], a fundamental challenge to the model arises from the fact that GWs from a binary neutron star merger [24] [25] traveled to the earth with the same speed (to within 1.7 s) as the electromagnetic

waves (EMWs) [47] [48]. This agreement implies, according to the model, that GWs have the same dispersionless index of refraction as EMWs when traveling through dark energy. Such a relation, as well as the proposed reduction in speed itself, falls well outside present general relativistic and electromagnetic theories, and supports the need for finding a unified theory of gravitation and electromagnetism, despite well-known failures in the past. Such a theory would also lead to an improved understanding of dark matter and dark energy.

Acknowledgements

I would like to thank Prof. George Fuller for his long-term interest in the work, and helpful suggestions. I am further indebted to Prof. Thomas Murphy for critical comments.

Conflicts of Interest

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

References

[1] Tangherlini, F.R. (2018) *Journal of Modern Physics*, **9**, 1827-1837. https://doi.org/10.4236/jmp.2018.99116

[2] Ade, P.A.R., *et al.* (2016) *Planck Collaboration*: *Astronomy and Astrophysics*, **594**, A18.

[3] Shadab, A., *et al.* (2016) *BOSS Collaboration*: *Monthly Notices of the Royal Astronomical Society*, 1-38.

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

[5] Schmidt, B., *et al.* (1998) *Astrophysical Journal*, **507**, 46-63.

[6] Perlmutter, S., *et al.* (1999) *Astrophysical Journal*, **517**, 565-586.

[7] Tangherlini, F.R. (2015) *Journal of Modern Physics*, **6**, 78-87. https://doi.org/10.4236/jmp.2015.61010

[8] Tangherlini, F.R. (2015) *Journal of Modern Physics*, **6**, 1360-1370. https://doi.org/10.4236/jmp.2015.69141

[9] Tangherlini, F.R. (2016) *Journal of Modern Physics*, **7**, 1829-1844. https://doi.org/10.4236/jmp.2016.713163

[10] Tangherlini, F.R. (2017) *Journal of Modern Physics*, **8**, 622-635. https://doi.org/10.4236/jmp.2017.84042

[11] Riess, A.G., *et al.* (2016) *SHoES Collaboration*: *Astrophysical Journal*, **826**, 56.

[12] Tangherlini, F.R. (2019) *Bulletin of the American Physical Society*, **64**, Abstract G1600007.

[13] Freedman, W.L., *et al.* (2019) *Astrophysical Journal*, **882**, 34. https://doi.org/10.3847/1538-4357/ab2f73

[14] Aartsen, M.G., *et al.* (2018) *IceCube Collaboration*: *Science*, **361**, 146.

[15] Riess, A.G., *et al.* (2001) *Astrophysical Journal*, **560**, 49-71.

[16] Chiba, T. and Nakamura, T. (1998) *Progress of Theoretical Physics*, **100**, 1077-1082.

https://doi.org/10.1143/PTP.100.1077

[17] Visser, M. (2004) *Classical Quantum Gravity*, **21**, 2603-2616. https://doi.org/10.1088/0264-9381/21/11/006

[18] Caldwell, R.R. and Kamionkowski, M. (2004) *Journal Cosmology Astroparticle Physics*, **9**, 009. https://doi.org/10.1088/1475-7516/2004/09/009

[19] Weinberg, S. (2008) Cosmology. Oxford University Press, New York, 32-33.

[20] Riess, A.G., Casertano, S., Yuan, W., Lucas, M. and Scornic, D. (2019) arxiv: 1903.07063

[21] Aghanim, N., *et al.* (2018) Planck Collaboration. arxiv: 1807.0609

[22] Addison, G.E., *et al.* (2018) *Astrophysical Journal*, **853**, 12. https://doi.org/10.3847/1538-4357/aaa1ed

[23] Macaulay, E., *et al.* (2019) *DES Collaboration*: *Monthly Notices of the Royal Astronomical Society*, **486**, 2184-2196.

[24] Ryan, J., Chen, Y. and Ratra, B. (2019) *Monthly Notices of the Royal Astronomical Society*, **488**, 3844-3856. https://doi.org/10.1093/mnras/stz1966

[25] Wong, K.C., *et al.* (2019) *HOLiCOW XIII*: *Monthly Notices of the Royal Astronomical Society*, 1-77.

[26] Abbot, B.P., *et al.* (2017) *LIGO Scientific Collaboration and Virgo Collaboration*: *Astrophysical Journal Letters*, **848**, L12.

[27] Abbot, B.P., *et al.* (2017) *Scientific Collaboration and Virgo Collaboration*: *Physical Review Letters*, **119**, Article ID: 161101.

[28] Abbot, B.P., *et al.* (2017) *Nature*, **551**, 85-88.

[29] Hotokezaka, K., *et al.* (2019) *Nature Astronomy*, **3**, 940-944. https://doi.org/10.1038/s41550-019-0820-1

[30] Chan, H.-Y., Fishbach, M. and Holz, D.E. (2018) *Nature*, **562**, 545-547. https://doi.org/10.1038/s41586-018-0606-0

[31] Poulin, V., Smith, T.L., Karwal, T. and Kamionkowski, M.G. (2019) *Physical Review Letters*, **122**, Article ID: 221301. https://doi.org/10.1103/PhysRevLett.122.221301

[32] Agrawal, P., Cyr-Racine, F.-Y., Pinner, D. and Randall, L. (2019) Rock "n" Roll Solutions to the Hubble Tension. arxiv: 1904.01016v1

[33] Blinov, N., Kelly, K., Kmjaic, G. and McDermott, S.D. (2019) *Physical Review Letters*, **123**, Article ID: 191102. https://doi.org/10.1103/PhysRevLett.123.191102

[34] Alexander, S. and McDonough, E. (2019) *Physics Letters B*, **797**, Article ID: 134830. https://doi.org/10.1016/j.physletb.2019.134830

[35] Hirata, K., *et al.* (1987) *Physical Review Letters*, **58**, 1490-1493. https://doi.org/10.1103/PhysRevLett.58.1490

[36] Bionta, R.M., *et al.* (1987) *Physical Review Letters*, **58**, 1494-1496.

[37] Alekseev, E.N., *et al.* (1987) *JETP Letters*, **45**, 589-592.

[38] Zwicky, F. (1933) *Hevitica Physica Acta*, **6**, 110-127.

[39] Tangherlini, F.R. (2018) *Journal of Modern Physics*, **9**, 573-583. https://doi.org/10.4236/jmp.2018.94039

[40] Aartsen, M.G., *et al.* (2018) *IceCube Collaboration*: *Science*, **361**, 147-156.

[41] Ackermann, M., *et al.* (2015) *Astrophysical Journal*, **810**.

[42] Krauss, L.M. and Chaboyer, B. (2003) *Science*, **299**, 65-69. https://doi.org/10.1126/science.1075631

[43] Gallart, C., *et al.* (2019) *Nature Astronomy*, **3**, 932-939. https://doi.org/10.1038/s41550-019-0829-5

[44] Gaia Collaboration, Gaia Data Release 2 (2018) *Astronomy and Astrophysics*, **616**, A1. https://doi.org/10.1051/0004-6361/201833955

[45] Lang, K.R. (1980) Astrophysical Formula. 2nd Edition, Springer, New York, 243.

[46] Aiola, S., *et al.* (2020) Atacama Cosmology Telescope: DR4 Maps and Cosmological Results. arxiv: 2007.07288

[47] Goldstein, A., *et al.* (2017) *Astrophysical Journal Letters*, **814**, L14.

[48] Savchenko, V., *et al.* (2017) *Astrophysical Journal Letters*, **848**, L15. https://doi.org/10.3847/2041-8213/aa8f94

Journal of Modern Physics, 2020, 11, 1236-1244
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Ring-Cavity Diode Laser

Jianing Han*, Lindsay Hutcherson, Kaori Munekane, Steven Shettlesworth

Physics Department, University of South Alabama, Mobile, AL, USA
Email: *jhan@southalabama.edu

How to cite this paper: Han, J.N., Hutcherson, L., Munekane, K. and Shettlesworth, S. (2020) A Ring-Cavity Diode Laser. *Journal of Modern Physics*, **11**, 1236-1244.

https://doi.org/10.4236/jmp.2020.118077

Received: October 17, 2019
Accepted: August 22, 2020
Published: August 25, 2020

Abstract

In this article, we report on a ring-cavity diode laser configuration, which has more features compared to the standing-wave external-cavity diode lasers. First, a transmission grating will be used, and the ring-cavity laser output will be the zero-order output from the transmission grating. Therefore, unlike the standing-wave external cavity diode lasers, the output direction of this ring-cavity laser does not depend on the laser's frequency. Second, since it is a ring-cavity laser, it can avoid the local heating caused by the antinodes of the standing-wave lasers. Moreover, because this ring-cavity will allow the laser beam to pass the transmission grating two times, it is expected that this will ultimately result in a narrower linewidth compared to the external-cavity standing-wave lasers with the same cavity length.

Keywords

PACS Numbers: 07.05.Fb, 42.15.Eq, 42.55.Px

1. Introduction

Van der Waals interactions are commonly known as attractive interactions, which is only true for ground-state atoms. Recent studies [1–4] show that repulsive van der Waals interactions can exist in excited states. Therefore, Rydberg atoms are essential for doing repulsive van der Waals interactions. The Rydberg excitation is often done by lasers, so lasers [5–7] are the key tools to study such interactions. Compared to dye lasers, an external-cavity diode laser has a much narrower bandwidth (about 1 MHz), and recent experiments are done by external-cavity diode lasers [1–4]. In this article, we will focus on external-cavity diode lasers. The cavity of diode lasers is made of the front and back surface of the diode. To improve the tunability of this type of laser and achieve a narrower linewidth, an external-cavity is often installed (Ref. [8] and references therein).

In this article, we report on a ring-cavity diode laser. This experiment is motivated by the following three facts: first, the widely used standing-wave external-cavity lasers [8,9] form standing waves between

a reflective grating and the back surface of the laser diode. In other words, a reflective grating and the back surface of the laser diode form the external cavity of the standing-wave diode lasers. The first order of the reflective grating is directly sent back to the laser diode, and the zero-order of this reflective grating is the output of the standing-wave external-cavity laser. Therefore, the laser's wavelength or frequency depends on the output beam's angle. In other words, once the laser's frequency is changed, the output direction will also change. If the laser frequency needs to be constantly adjusted, this will cause a lot of downstream alignment issues. This is an important consideration in Rydberg physics. The output of this ring-cavity laser proposed is the zero-order of a transmission grating instead of a reflective grating. Therefore, the output direction will not be affected when the frequency is changed, or the output direction does not depend on the laser's frequency. Similar to the Littman configurations [10, 11], the configuration presented in this article applies to diode lasers. This is crucial for studying repulsive van der Waals interactions, which only exist in excited states [1–4]. Second, this type of laser is expected to produce a narrower linewidth compared to the standing-wave external-cavity diode lasers. The linewidth of the ring-cavity laser can be reduced by allowing the laser beam within the cavity to pass the transmission grating two times. Third, the ring-cavity laser [12] can avoid the local heating caused by the antinodes of the standing waves. A standing wave is formed by two counter-propagating waves. Within a standing-wave cavity, there are nodes and anti-nodes along the cavity, or the intensity variation, which will result in non-uniform heating, or local heating, of the laser medium or a diode in this case. The ring-cavity configuration can avoid local heating [13]. The waves in a ring-cavity are propagating waves, so no nodes or antinodes will be formed; the heating will be uniform. The laser intensity stability of the standing-wave external-cavity lasers is partially caused by the local heating from the antinodes of the standing wave in the laser diode, especially at high intensities, which will limit the ability of creating bigger MOTs (Magneto-Optical Traps).

This ring-cavity laser has many potential applications. Related to earlier discussions, lasers can be used to excite atoms from ground states to excited states, or Rydberg states [14, 15]. In addition, the standing-wave external-cavity diode lasers [8, 9, 16] are the workhorses in the cold atom field. It can be used for laser cooling and trapping. Since the output is the zero-order output of a transmission grating, this will make the optics adjustment a lot easier. This system can be used to study the amplification of the ring-cavity laser using tapered amplifiers or other types of lasers [17]. In addition, this system can be used as a frequency comb for high precision measurement [18] and laser spectroscopy, etc.

This paper is arranged in the following way: the experiment is introduced in the next paragraph, which is followed by the tunability measurement. The output power from the ring-cavity laser is then discussed and analyzed. A conclusion is given at the end.

2. Experiment

In this experiment, we use a PLT5 488 diode, manufactured by OSRAM Opto Semiconductors, mounted in a temperature and current controller kit purchased from Thorlabs, which includes the laser mount. Figure 1 shows the schematic diagram of this ring-cavity laser. The

grey line is the laser beam pass. The beam coming out of the laser first passes through a grating. In order to avoid the zero-order back reflection from the grating, it is recommended to tilt the angle of the grating slightly, so the grating is not parallel to the laser diode surface. After the grating, the beam diffracts. The first order will be reflected by mirror 1, then reflected again by mirror 2. The reflected beam from mirror 2 is sent back to the laser along the opposite direction of the −1 order direction of the grating. In other words, the laser beam passes the grating again. As described earlier, this laser can produce a narrower linewidth than the standing wave lasers with the same cycle length by allowing the laser beam to pass the grating two times. Here the cycle length is the total length for the laser beam to pass one cycle. For example, the cycle length of a ring cavity is the same as the cavity length $(2l_1 + l_2 + l_3 + l_4)$, while the cycle length is $2l'$ for a standing wave cavity, where l' is the distance between the two standing wave cavity mirrors. After the final reflected beam enters the grating, the laser beam passes the laser diode, reaches the back surface of the diode, and completes the ring-cavity. The zero-order is the output of the ring-cavity laser. The laser beam is collimated to 0.6 mm in diameter, which is measured at I = 90 mA using the knife-edge method, a knife blade mounted on a translation stage. The grating used in this particular experiment is a 1000 line per millimeter transmission grating. The first-order diffraction efficiency, (the output power in the first-order diffracted beam)/(the output power without the grating), is about 35.4%, and the second-order diffraction efficiency is about 5.33%. The third and above order diffraction efficiencies are zero. Both diffraction efficiencies are measured at 90 mA, and the total output power of the diode before the grating at this current is about 52.5 mW. The mirrors are aluminum-coated mirrors. The reason that aluminum-coated mirrors are used is that those mirrors have higher reflectivity in the wavelength range that we studied and low transmissivity. The total cavity length of this ring-cavity, $2l_1 + l_2 + l_3 + l_4$, is about 16 cm. The linewidth measurement was done through a Fabry-Perot interferometer (model number: SA200-3B) borrowed from Thorlabs. The free spectrum range of this Fabry-Perot interferometer is 1.5 GHz, and the resolution is 7.5 MHz. We connect the Fabry-Perot interferometer output to an oscilloscope. The oscilloscope is then connected to a GPIB board, which is connected with a computer. All the oscilloscope traces are recorded on a computer through this GPIB board.

3. Result and Discussion

The Q factor of this cavity can be calculated by the following equation [5]:

$$Q = \frac{v_0}{v_F} F, \tag{1}$$

where v_0 is the resonant frequency, $v_0 = \frac{c}{488nm}$ in our case, and c is the speed of light. v_F is the cavity frequency difference, $v_F = \frac{c}{2l_1+l_2+l_3+l_4}$ in this cavity (Figure 1). In addition, F is the Finesse of the cavity shown in Figure 1:

$$F = \frac{\pi r \sqrt{r_1}}{1 - r_1 r^2}, \tag{2}$$

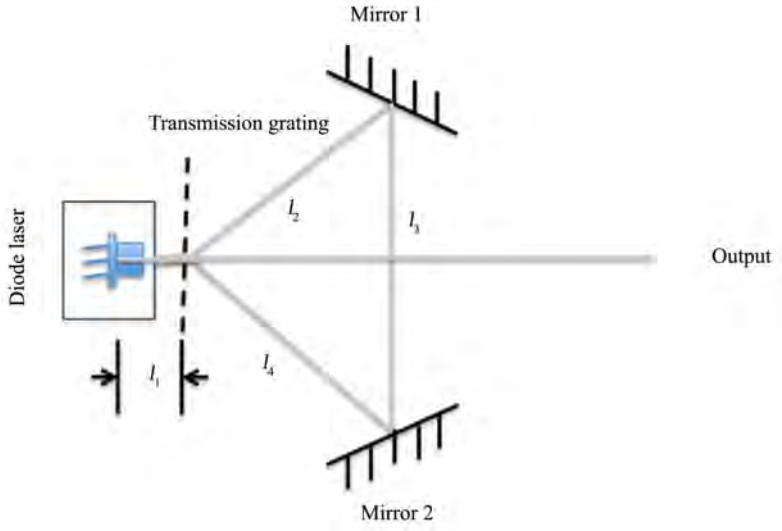


Figure 1. The schematic diagram of the ring-cavity laser. l_1 is the distance between the back surface of the diode laser and the grating. l_2 is the distance between the grating and mirror 1. l_3 is the distance between mirror 1 and mirror 2. l_4 is the distance between mirror 2 and the grating.

where r_1 and r are the amplitude attenuation factor introduced by the reflection from the back side of the diode and the reflection of one of the mirrors. The absorption or energy loss by the grating and the medium is also included in these terms. r_1 is much smaller than r in our case, since r_1 includes the energy loss caused by the reflection from the front side of the diode. If $r \approx 80\%$ and $r_1 \approx 50\%$, the Q factor of this 488 nm external laser cavity shown in Figure 1 is about 8.6×10^5.

One of the features of this ring-cavity laser is that the frequency is tunable. The way to understand this is that the cavity selects a certain frequency, which is used as a seed to be sent back to the cavity. This seed beam will pull the frequency of the zero-order beam to this selected frequency. Therefore, by changing the cavity length, the laser frequency can be changed. To show the tunability, we first show that we are able to measure the laser's relative output frequency. We send the ring cavity laser output to a Fabry-Perot interferometer. Figure 2 shows two oscilloscope traces taken from the Fabry-Perot interferometer. To reduce the power broadening and laser mode hopping at high intensity, we set the laser diode current at a low current. The laser current in Figure 2 and Figure 3 is below 55 mA. Both traces are taken by directly connecting the photodiode of the Fabry-Perot interferometer to an oscilloscope. A time-dependent voltage controlling the Fabry-Perot cavity length is added to the Fabry-Perot cavity. As the voltage changes, the light intensity as a function of time collected by the photodiode is sent to the oscilloscope. The left figure, Figure 2(a), is a long scan. The horizontal axis is converted from the original time to frequency by setting the frequency difference between the two peaks as 1.5 GHz, the free-spectrum range of this Fabry-Perot interferometer. The off-set frequency can be calculated by

$$f = \frac{c}{\lambda}, \tag{3}$$

where c is the speed of light, and $\lambda = 488$ nm is the wavelength of this laser. The off-set frequency is calculated to be about 6.15×10^5 GHz.

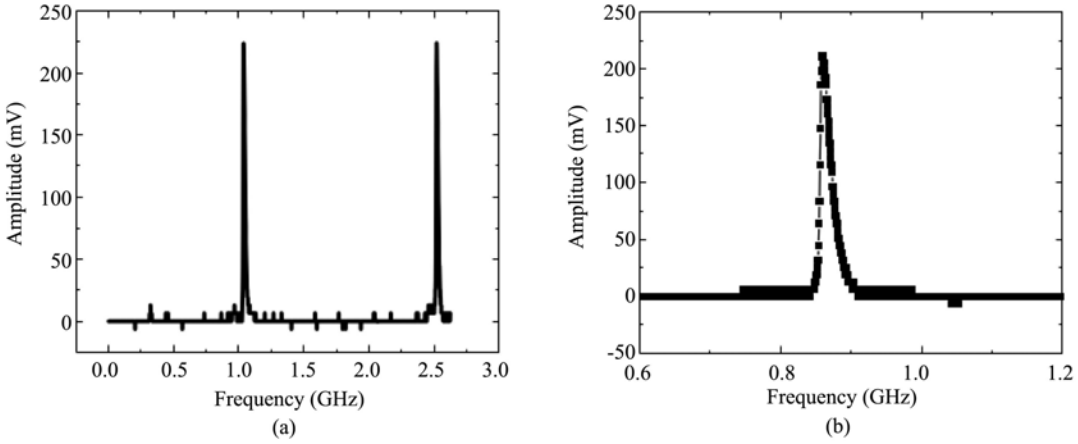


Figure 2. (a) is a long scan of the trace; (b) The detail of a single peak.

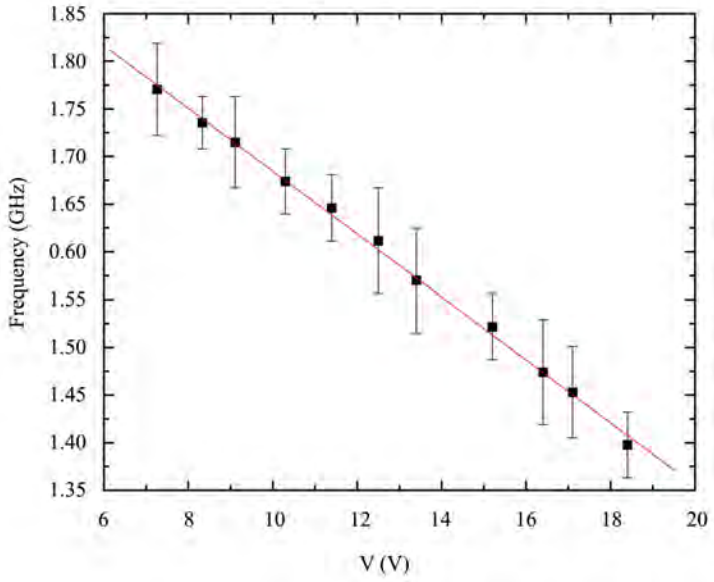


Figure 3. The relative frequency of the laser as a function of the voltage on the piezoelectric actuator. The red solid line is the linear fit of the experimental data.

To show the details of one of the peaks, we expand the horizontal axis of the oscilloscope and collect a trace for one of the peaks, as shown in Figure 2(b). The FWHM (full width at half maximum) of the trace shown in Figure 2(b) is about 17 MHz, which is close to the resolution of this Fabry-Perot interferometer.

The frequency of this type of laser is adjustable. To change the output frequency of this ring-cavity laser, we change the angle of mirror 1 or 2. In this experiment, we use a piezoelectric actuator (PZT) to change the angle of mirror 2. Alternatively, we can change the angle of mirror 1, and similar results are expected. In this experiment, the piezoelectric actuator is installed in the mirror mount of mirror 2 to control the horizontal direction of the laser beam reflected off of mirror 2. Figure 3 shows the relative laser frequency as a function of the voltage on the piezoelectric actuator, which is used to change the horizontal angle of mirror 2. As we vary the voltage on the piezoelectric actuator, the laser output frequency changes as a function of the

voltage on the piezoelectric actuator. The data points are obtained in the following way: at each angle or each voltage on the piezoelectric actuator, a few oscilloscope traces are taken. Due to the thermal fluctuation and mechanical vibration of the laser mount, the frequency of the laser jitters. In addition, the power from the laser fluctuates, which will cause the uncertainty of the linewidth of the laser. Therefore, at each voltage, we find the minimum frequency of the few traces taken at a particular voltage and the maximum frequency of those few traces, instead of calculating the average of few frequencies measured at one voltage. We then calculate the average of those two, $\frac{maximum+minimum}{2}$, to get the actual data point shown in Figure 3 as the black solid squares. The error bar is calculated by the difference between the maximum frequency and the average frequency or the difference between the average frequency and the minimum frequency. In other words, the total length of the error bar is (maximum-minimum). The total frequency tuning range of this laser is 488 nm ± 1 nm. The free-running laser linewidth is greater than 1.5 GHz.

In addition to the tunability measurement, we also studied the power of this ring-cavity laser. Figure 4 shows the output power comparison between the free-running laser, plotted as red dots, and the output power from the ring-cavity laser, plotted as black squares. The power is measured by a thermopile power meter. This graph shows that the power increases as the current increases, which is what we expected. Comparing those two sets of data, it is also shown that the output power of this ring-cavity laser is much lower than the free-running laser. Specifically, the ring-cavity laser output is only about 20% of the free-running laser output. In this experiment, since we allow the laser beam to pass the grating two times, we would expect the loss is about twice the power loss from the standing-wave external-cavity laser [8, 9]. Here we again use the 1000 line/mm grating, and both the first order and second order from the grating are visible, which means that for every ring cycle, the second-order output will cause power loss. This can be avoided by choosing a grating with a higher line density, or the line density greater than 1000 lines per millimeter for a 488 nm laser. This is something that we would like to explore in the near future. The error bars are based on the maximum power fluctuation from the power meter reading. During the measurement, we noticed that the power typically drifts downward once the current sent to the diode laser is fixed.

We have also tested the mirrors' reflectivity effect on this ring-cavity laser. We changed the reflectivity of the mirrors from 32% to 80%, and the output difference is less than 3%. In other words, it turns out that by increasing the reflectivity of the cavity mirror, the effect of the mirrors' reflectivity on the output power of the ring-cavity laser is negligible.

Within the ring-cavity laser, we had to increase the current limit of the temperature controller, which controls the temperature of the laser diode, in order to stabilize the temperature of the ring-cavity laser unit. This wasn't the case with the free-running laser. In other words, the free-running laser does not require high current to be stabilized, which indicates that the free-running laser generates less heat than the ring-cavity laser. This may be related to the uniform heating of the laser system or caused by the high intensity within the laser cavity.

Here are a few limitations of this ring-cavity laser. First, the output power is very low. Choosing a grating, which has a greater line

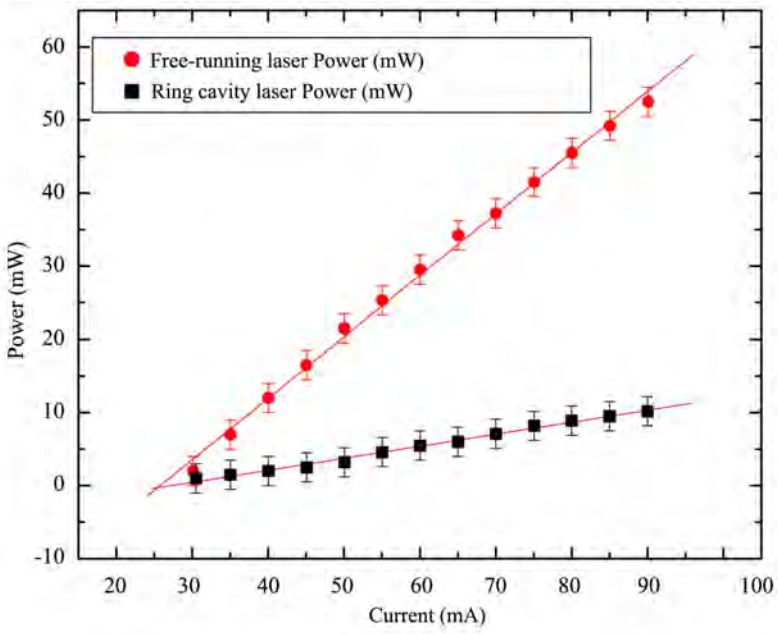


Figure 4. The output power from the free-running laser (•) and the output from the ring-cavity laser (■). The red solid lines are the linear fits of the experimental data.

density, will be helpful. One limitation of diode lasers is the mode-hopping. Very similar to other diode lasers, this ring-cavity laser also shows mode-hopping, which may be improved by stabilizing the optical table and more stable temperature and current control systems. It is generally observed that as the current increases, the mode hopping rate increases, and the probability of multimode also increases. Changing the internal and external cavity length may help to reduce the mode-hopping rate.

4. Conclusion

In conclusion, we have shown a ring-cavity diode laser that is tunable. The frequency can be easily adjusted by adjusting the voltage of a piezoelectric actuator installed in one of the mirror mounts. In addition, this laser output is the zero-order output of a transmission grating; therefore, the direction of this output is frequency independent. This type of laser is very useful for studying excited states atoms, especially for repulsive van der Waals interactions. Moreover, this type of laser is promising to produce narrower linewidth than the standing-wave external-cavity laser with the same cavity length by allowing the laser beam to pass the grating two times.

Acknowledgements

The authors would like to thank the equipment loan from Thorlabs. This material is based upon work supported by the U. S. Army Research Office (ARO), the travel support from DOE/EPSCOR, and the Air Force Office of Scientific Research (AFOSR). It is our pleasure to acknowledge Dr. Xiangdong Zhang for writing the program for data collection.

Conflicts of Interest

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

References

[1] Han, J. and Gallagher, T.F. (2009) *Physical Review A*, **77**, Article ID: 015404.

[2] Han, J. and Hu, C.Y. (2016) *Molecular Physics*, **114**, 637. https://doi.org/10.1080/00268976.2015.1109150

[3] Han, J. (2017) *Physical Review A*, **95**, Article ID: 062502.

[4] Han, J. (2017) *Molecular Physics*, **115**, 2479.

[5] Bhaa, S.E.A. and Malvin, T.C. (1991) Fundamentals of Photonics. John Wiley & Sons Inc., New York.

[6] Zhou, B.K. (1980) Laser Theory. National Defense Industry Press, Beijing.

[7] Wyatt, R. and Devlin, W.J. (1983) *Electronic Letters*, **19**, 110-112. https://doi.org/10.1049/el:19830079

[8] Wieman, C.E. and Hollberg, L. (1991) *Review of Scientific Instruments*, **62**, 1-20. https://doi.org/10.1063/1.1142305

[9] MacAdam, K.B., Steinbach, A. and Wieman, C. (1992) *American Journal of Physics*, **60**, 1098-1111. https://doi.org/10.1119/1.16955

[10] Littman, M.G. and Metcalf, H.J. (1978) *Applied Optics*, **17**, 2224-2227. https://doi.org/10.1364/AO.17.002224

[11] Littman, M.G. (1978) *Optics Letters*, **3**, 138-140. https://doi.org/10.1364/OL.3.000138

[12] Shih, M.C., Wang, S.C. and Liang, C.W. (2005) *Journal of Vacuum Science & Technology B*, **23**, 2180-2184. https://doi.org/10.1116/1.2009772

[13] Soletsky, P.A., Bixler, D.L. and Dunning, F.B. (1994) *Review of Scientific Instruments*, **65**, 3585-3587. https://doi.org/10.1063/1.1144541

[14] Gallagher, T.F. (1994) Rydberg Atoms. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511524530

[15] Demtroder, W. (2003) Laser Spectroscopy. 3rd Edition, Springer-Verlag, Berlin, Heidelberg, New York. https://doi.org/10.1007/978-3-662-05155-9

[16] Matthew, S.C. (2003) External Cavity Diode Lasers: Controlling Laser Output via Optical Feedback. B.S. Thesis, College of William and Mary, Williamsburg, VA.

[17] Ryan, C. (2015) Construction and Optimization of a Tapered Amplifier System for Applications in Ultra-Cold Plasma Research. Honors Thesis (Adviser Duncan Tate), Colby College, Waterville, ME.

[18] Jun, Y. and Steven, C.T. (2005) Femtosecond Optical Frequency Comb: Principle, Operation, and Application. Springer, New York.

Journal of Modern Physics, 2020, 11, 1245-1261
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

The Experiments Detecting of Real Magnetic Charges in Structures of Atoms and Substance

Robert A. Sizov*

Solid State Physics Institute, Moscow, Russia
Email: sizov.robert@gmail.com

How to cite this paper: Sizov, R.A. (2020) The Experiments Detecting of Real Magnetic Charges in Structures of Atoms and Substance. *Journal of Modern Physics*, **11**, 1245-1261.
https://doi.org/10.4236/jmp.2020.118078

Received: July 19, 2020
Accepted: August 22, 2020
Published: August 25, 2020

Abstract

Magnetic neutron scattering in Y-type hexagonal ferrite crystals, studied by the author in 1968-1971 and presented in the article showed that the entire density of the so-called magnetic moments of Fe^{3+} ions can significantly shift from the position of their nuclei. As result of these shift the structure in form of the chain magnetic spiral is realized in ferrite lattice. The noted shifts of the "magnetic moments" served as the basis for the author's assumption that these "moments" are "fig sheets" behind which the magnetic poles (magnetic charges) real existing in the shells of atoms are hidden. In this case, the scattering of neutrons is carried out by magnetic charges, and not theoretical surrogates in the form of magnetic moments. In addition to participating in atomic structures, magnetic charges populate potential conduction zones in conductors, where they are exist in compositions of magnetic dipoles. Under the influence of an external magnetic field, a polarization of magnetic dipoles is realized in the conductor, the field strengths of which are directed against the external magnetic field. It is these dipole magnetic fields that are responsible for such a well-known physical phenomenon as diamagnetism. Under the conditions of noted polarization of magnetic dipoles the author managed to perform mechanical separation of magnetic charges in pairs ±g and to charge experienced bodies (metal plates) by the magnetic charges of one sign. The fact of such a charging was detected through magnetostatic interaction between magnetic charges on the plates using highly sensitive torsion balances. This experiment is presented in detail in this article. The results of these experiments, as well as subsequent experimental and theoretical studies of the author, which, in general composition, were carried out from 1968 to the present, showed that magnetic charges are real structural components of the atoms and substance. So, for example, the atomic shells are not electronic but electromagnetic. The main reason that real magnetic charges were

*Author is the doctor of physical and mathematical sciences. At present works as individual researcher.

"buried alive" in the existing physical theories is the physics of their confinement in substance forces of which, in its rigidity, is many times greater than the electron confinement forces.

Keywords

Magnetic Charges, Magnetons, Antimagnetons, True Antielectrons, Magnetization, Magnetic Structures, Torsion Balance, The Magnetostatic Interaction

1. Introduction

The first person to experimentally observe real magnetic charges in a substance was the remarkable Austrian physicist Felix Ehrenhaft. Based on the results of his research which lasted for more than 40 years, he published about 30 articles in physical journals [1]. It is important to note that Ehrenhaft's experiments were repeated by numerous followers who confirmed his results and conclusions (see, for example, [2]).

The Ehrenhaft experiments are a magnetic analog of well-known Milliken tests from determining the value of the electron charge. Very small particles of solids substance were placed in a vertical uniform magnetic field free of residual electrical charges. The particles were illuminated by concentrated beam of light. The optical system allowed determining the parameters of the particles movement. The basic experimental result of investigation Ehrenhaft consisted in the fact that it was found a logical movement of particles along power lines of the magnetic field. With the change in the direction of the field, the direction of movement of the particles also was changed. According to conclusions Ehrenhaft motions of particle, that observed in his experiments, are determined their charging by the magnetic charges of different signs.

However, interpretation of the experimental results which was built by Ehrenhaft on analogy with behavior of electric particles in the electrostatic field, does not appear convincing enough. Besides, values of forces in the observed interactions were comparable with parasitic, so called, radiometric forces. It is also obvious and what such serious conclusions which were made by Ehrenhaft, has need in deepening and first of all in of development of conceptions about physical parameters of magnetic charges (magnetic spinor particles) and their place within the substance structures. One way or another, but the experiments of Ehrenhaft which showed interesting and very important results, did not receive recognition and in the present were practically forgotten.

In 1968-1971 when conducting a study of the magnetic scattering of neutrons in ferrites with a hexagonal structure of type Y and Z at the TWR ITEP reactor in Moscow, the author discovered and investigated a very unusual phenomenon, which, according to the established model version, consisted in a density shift of

so-called magnetic moments of Fe^{3+} ions from the position of their nuclei over very considerable distance (up to ~0.9 Å) [3] [4]. It is the noted displacements that served as the basis for the author's assumption about the possible existence of real magnetic poles (magnetic charges) in the structures of atoms and substance. The results of the studies and conclusions of Ehrenhaft which the author discovered later, greatly contributed to strengthening his confidence in the correctness of initial assumption.

Subsequent experiments of the author which, with interruptions, continued until 1992, allowed him, with a high degree of probability to assert, that magnetic charges are real structural components of atoms and substance. In total, over the 24 years of his experimental works (1968-1992) the author performed 8 experiments of which 6 showed a positive result and 2 experiments were negative. The results of the experiments were evaluated based on their compliance with the author's concept: magnetic charges are real structural components of atoms and substance. These results are very original and important information about the real structure of all types of physical mass (atoms, nucleons, substance and others).All these experiments, since 2001, are described in detail in the author's Russian-language books [5] [6]. In addition, the results of his four basic experiments are briefly presented in the English-language article [7].

However, the existing information regarding by the author's studies of real magnetic charges was insufficient for a wide physical audience. This circumstance was indicated to the author by scientists interested in this problem. Therefore, work was undertaken on an expanded presentation the results of his experimental studies of magnetic charges in English format.

The purpose of this study is the experimental detection of real magnetic charges in matter, as well as the determination of their physical parameters.

2. Magnetic Neutron Scattering in Structure of the Hexagonal Ferrite Type Y

When analyzing the results of the neutron diffraction in Y-type hexagonal ferrite crystals with the chemical formula $Ba_{0.4}Sr_{1.6}Zn_2Fe_{12}O_{22}$, for to satisfy be the experimental data, the author had to shift the entire density so-called magnetic moments (hereinafter, the "magnetic moments" or "MM") of Fe^{3+} ions from the positions of its nuclei, as shown in **Figure 1**. According to the accepted model, which was formed in result of the analysis the intensity and location on the neutronograms of more than 30 magnetic reflexes of the 00l and h0l series, the chain spiral-block magnetic structure is realized in the crystal lattice investigated ferrite.

The feature of such a magnetic structure is that the helical ordering is formed not by individual "magnetic moments" of Fe^{3+} ions, but by their groups in a part (block) of the ferrite unit cell which shown in **Figure 1**. Inside each block, the "magnetic moments" correspond to collinear ordering, that is, so as shown in **Figure 1**, but between the blocks, when moving along the Z axis, is the turn of "MM" by constant angle is realized.

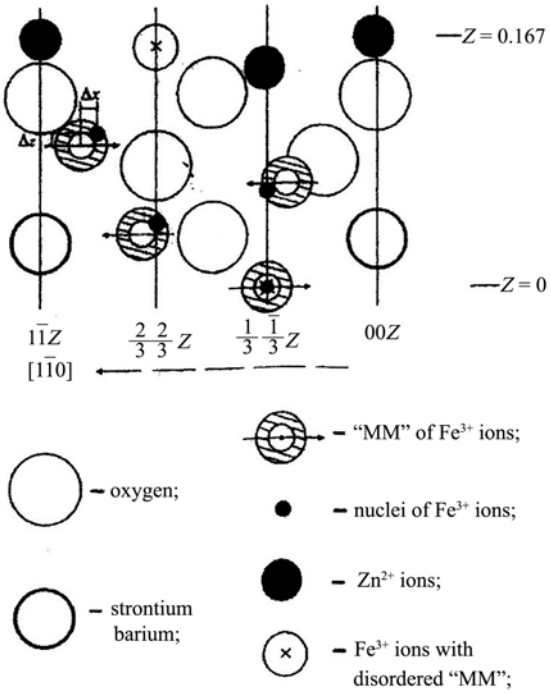


Figure 1. A model of the magnetic structure with a shift in the density of atomic "magnetic moments" of Fe^{3+} ions from their nuclei in the composition of "collinear" unit cell block in type Y hexagonal ferrite. The atomic "magnetic moments" densities responsible for magnetic neutron scattering are marked in **Figure 1** hatch. The small black circles indicate the nuclei of ions located in the center of distribution of their electronic (according to the author, electromagnetic) density.

In this case, the period of the magnetic spiral, the axis of which is directed along the hexagonal axis of the unit cell of ferrite, turned out to be a multiple of the block size along the z axis.

The model of the magnetic structure with a density shift of "MM", presented in **Figure 1**, best suited neutron diffraction data and not only in terms of the convergence of experimental and calculated values of the intensity of diffraction reflections but and allowed to index all the observed magnetic reflexes. It is important to note that the author's successful indexing of magnetic reflections of the h0l series is a self-sufficient evidence in favor of the model of the chain helical-block ordering of "magnetic moments" in the lattice of the studied ferrite. The fact is that the magnetic reflexes of the h0l series, in contrast to the 00*l* reflexes, were not detected as the paired satellite reflexes at nuclear reflections, which should have taken place when the magnetic density localized on ionic nuclei. In addition, based on the accepted model of the magnetic structure it was possible to explain, the best way, the anomalous intensity ratios in the pairs of magnetic satellites $00l^{\pm}$ which were observed at nuclear reflections of the 00l series.

The marked indication of all observed magnetic reflections, as well as a fairly positive R-factor, *i.e.* the convergence of the experimental and calculated values of their intensity (14%) gives give reason to make sure that the selected model is

correct. Thus, the "magnetic moments", shifting from the nuclei, are located along a spatial spiral with radii Δx located in the plane of the basis (00*l*) of the hexagonal unit cell. The author established the magnitudes and directions of such displacements for various crystallographic positions of iron ions in the Y-type ferrite lattice, the maximum values of which were $\Delta x_{max} = 0.7$ Å and $\Delta z_{max} = 0.4$ Å with an error of ±0.1 Å.

It is also important to note that subsequent x-ray diffraction studies of the same ferrite crystals, conducted at the Institute Crystallography of the Academy of Sciences in Moscow, did not reveal any delocalization of electron density. As know, X-ray quanta are scattered on the so-called compensated electron density of atomic shells (according to the author's research this density is electromagnetic) and do not "feel" the uncompensated charged density which and is of the density of magnetic charges. At the same time, X-ray diffraction studies proved to be very useful, since they allowed us to separate electron and magnetic charge densities in our representations of the Fe^{3+} ion shell. In other words, the charge density responsible for the magnetic scattering of neutrons is the density of real magnetic charges. As for electrons in atomic shells, they exist in the composition of the so-called compensated (electromagnetic) charged density, which is the source of the vortex electromagnetic (gravitational) field. The latter density is capable of scattering x-ray (electromagnetic) quanta, but cannot scatter neutrons. Details of this neutron-structural study can be found in the original publication of the author, which is presented in the English language [8].

Subsequent experiments of the author which, with breaks, were carried out until 1992, confirmed his initial assumption and served as the basis for the final statement that magnetic spinor particles (fundamental magnetic charges) are real structural components of atoms and substance. As for the shells of atoms, they are electromagnetic, not electronic, and consist of electric and magnetic charges [9]. Hence and the quotation marks around the term "magnetic moments", which in existing physics are completely mistakenly associated with the so-called atomic 3d and 4f electrons.

Note 1. In the process of interpretation, as well as publication of the results of neutronographic studies of hexagonal ferrites, the author remained within the framework of the generally accepted concept of atomic magnetic moments. Therefore, these concepts are given in these articles without quotes. The idea that magnetic moments are "Fig leaves" hiding real magnetic charges came to the author later, after performing direct experiments to detect them.

3. Experiments on Charging Test Bodies by the Magnetic Charges of Various Signs. Magnetostatic Interaction of Bodies Charged of Magnetic Charges

3.1. The Technology of Charging (the Magnetization) of Test Bodies by Magnetic Charges of Different Signs

In contrast to the search for an electrified Dirac monopole [10], in this study as-

sumed that magnetic charges are fundamental spinor (charged) particles that actually exist in atomic structures and are magnetic analogues of known atomic electric particles [11]. For example, magnetic charges in atomic shells have charges and spins equal to the corresponding values for electrons. It is important to note that the main state of magnetic charges in Nature is their existence in the composition of magnetic dipoles. The processes of breaking the marked magnetic dipole bonds in a substance lead to its **magnetization**, what is a magnetic analog of the processes of the **electrization** substance. In addition to participating in the structures of atoms, magnetic charges in conductors populate both potential and in some special cases (superconductivity) real conduction zones.

When an external magnetic field is applied at substance, for example, at conductors, so polarization of the marked magnetic dipoles realized, as shown in Figure 2. In this case, the vector of the intrinsic dipole magnetic field **h** is directed against the external magnetic field (rot**H**) what is simple explanation for such a physical phenomenon as **diamagnetism**.

Note 2. In a similar way explains the occurrence of **giant diamagnetism** in superconductors *i.e.* **Meissner effect**. Under the conditions of the superconducting state, the number of magnetic charges (magnetic dipoles) free of communication with the lattice of superconductor increases by several orders. These dipoles are able, through their dipole fields **h**, to resist an external magnetic field (up to the value of $\mathbf{H}_{cr}$) even in the surface layer of a superconductor.

If process of the polarization of magnetic dipoles, the scheme of which is shown in Figure 2, materializes in practice, it is possible to carry out the charging of surface of test bodies with magnetic charges of different signs. For this, it is necessary, under the influence of an external magnetic field, to cut the test body along the FT plane (see Figure 2). In this case, with some probability, the separation of magnetic charges in the pairs ±g can occur.

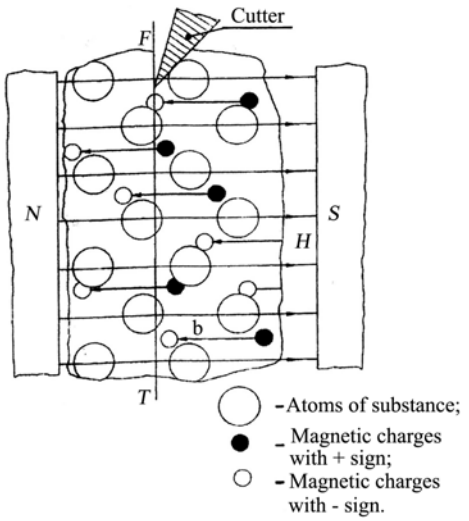


Figure 2. The scheme of the technology for the mechanical separation of magnetic charges in ±g pairs in the conditions of magnetic dipoles polarization in conductors under the influence of an external magnetic field.

As a result of the process described above the adjacent planes FT' and FT'' of the test body, after its mechanical separation in a magnetic field, can turn out to be charged magnetic charges of different signs. The signs of charges on the test bodies can be indicated relative to the poles of the magnet in which was process of dividing the body as shown in Figure 2. The fact that the surface of test bodies is charged with magnetic charges of different signs can be established using the effects of the magnetostatic interaction of charged samples.

The experimental technique of charging bodies with magnetic charges of different signs consisted in the fact that plates of the various metals were fixed at one of the poles of an electromagnet and in magnetic field of ~5 - 10 kOe, their surface layers were removed mechanically, manually, using a cutting tool. The cutter was a stainless steel plate measuring 3.5 × 30 cm^2 and 4 mm thick. One end of this plate was rounded with an approximate radius of 6 - 7 cm, so that its sharp edges were formed. Using these sharp edges, the layers were scraped off the test bodies (the plates: copper, nickel, lead, cadmium) in a magnetic field. The plates were mounted on the magnet pole by means of glue (epoxy resin).

Note 3. It should be noted that the author was aware that the possible magnitudes of the magnetic charges on test bodies, which can be realized within the framework of the described technology, will turn out to be very small. Therefore, it was necessary to develop and use experimental equipment that would ensure the sensitivity of the forces interaction between magnetic charges on the test bodies in the range of 10^{-5} - 10^{-8} dyne.

3.2. Magnetostatic Interaction of Test Bodies Charged with Magnetic Charges of Different Signs

The study of magnetostatic interaction of magnetic charges on a plates, after the separation of charges in pairs ±g in the magnetic field described above, was carried out in vacuum (10^{-3} - 10^{-4} атм.) at 20˚C using the highly sensitive torsion scales (Coulomb scales) shown in Figure 3.

The scales consisted of a body 1 and pipe 2 made of stainless steel. In the upper part of the tube, an annealed tungsten filament 4 with a diameter of 20 microns and a length of 50 cm was suspended on a holder 3. To the lower end of the filament was attached a rocker 5 made of aluminum, with a shoulder length of 10 cm. At the ends of the rocker arm were fixed: a test plate charged with magnetic charges 6 and an uncharged counterweight plate 7, placed in the gap of the damper (compensator) vibrations of the rocker arm, of the condenser type 8. On two rotating holders 9, located to the left and right of the test plate on the rocker, three test plates 10 were fixed (two charged with magnetic charges of different signs and one uncharged). The dimensions of the plates were 2 × 2 cm^2 with a thickness of 1 mm. Rotating the axis of the holders 11, it was possible to carry out a variety of magnetostatic interactions using two or three interacting plates. The mirror 13 was mounted on vertical part of the rocker arm, which together with a laser source and a screen-scale 3.1 m apart from the mirror, made up a system for detecting the movement of a movable plate on the rocker arm in

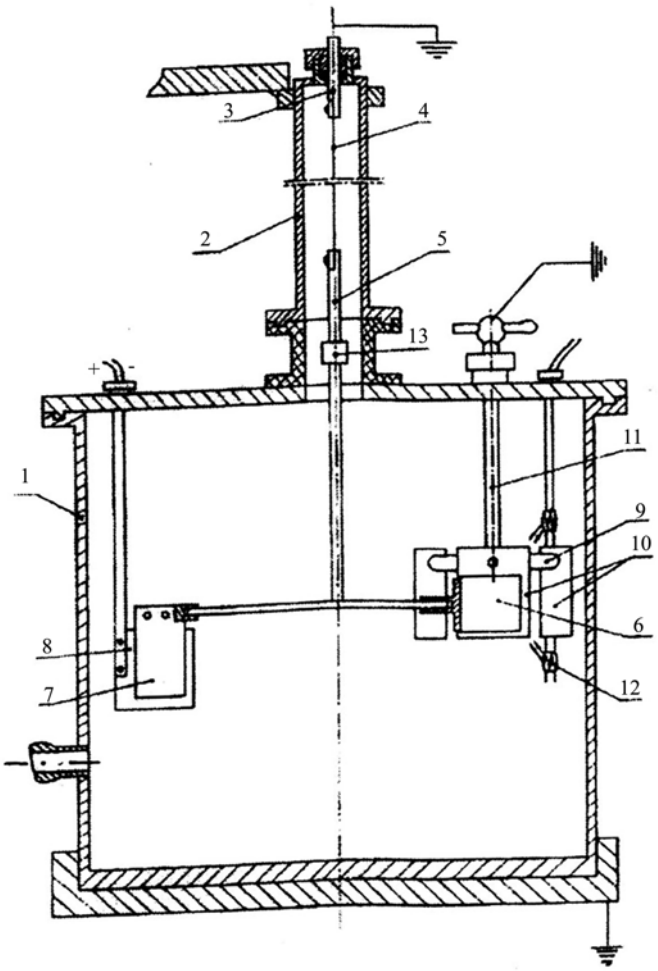


Figure 3. The schematic diagram device of torsion scales which used to detect the magnetostatic interaction of bodies charged with magnetic charges.

the studied interactions. The maximum possible deviation of the light spot on the scale-screen, under the conditions of the geometry of the present experiment, was ~20 cm. The experiment also used the tungsten light emitters 12, which allow one to study the effect of photons on the state of magnetic charge on of the test plates. The emitters were two short ceramic tubes with a winding of tungsten wire, which was powered by alternating electric current with a voltage of 220 V. The tubes in the balance housing were attached by means of a metal holder (not shown in Figure 3).

Stagestechnologyofexperiment.

1) The bringing the measuring system to the zero position under conditions when the experimental plates on the holders 9 are allotted to the maximum distance from the plate on the rocker;

2) The determining the position of the deflected light spot on the scale-screen as a result of magnetostatic interaction of magnetic charges on the plates after calming of oscillations a rocker by means of compensator 8.

3) The bringing the measuring system to the zero position, when the plates on

the holders 9 remote to the maximum distance from the plate on the rocker arm.

The main result of each measurement was the sign of the deviation of the light pulse reflected from the mirror on the scale-screen relative to the zero position of the registration system. For of interacting plates, at the maximum, 25 counts were performed which amounted to a complete measurement cycle for the selected set of interacting charges on the plates. In addition, as a finished result within the framework as part of our experiments, the condition was accepted when the number of positive metering by 5 units exceeded the number of negative.

To eliminate electrostatic interference, all parts of the balance were grounded using good "ground". In addition, the test plates, as well as the details of the rocker arm, were demagnetized in a solenoid powered by an alternating current of 50 hertz.

3.3. The Experimental Results

The results of magnetostatic interaction of plates charged with magnetic charges are shown in **Figure 4** (copper plates) and **Figure 5** (nickel plates) for different sets of magnetic charge signs on interacting plates. Sets of magnetic charges on the plates involved in the experiments are shown in the upper part of **Figure 4** and **Figure 5**. **Figure 4(a)** and **Figure 5(a)** show the results of interactions for the case when the movable plate fixed on the rocker was charged with magnetic charges with a plus sign, and **Figure 4(b)** and **Figure 5(b)** of the charged with charges with a minus sign.

The results of magnetostatic interaction of plates charged with magnetic charges are shown in **Figure 4** (copper plates) and **Figure 5** (nickel plates) for different sets of magnetic charge signs on interacting plates. Sets of magnetic charges on the plates involved in the experiments are shown in the upper part of **Figure 4** and **Figure 5**. **Figure 4(a)** and **Figure 5(a)** show the results of interactions for the case when the movable plate fixed on the rocker was charged with magnetic charges with a plus sign, and **Figure 4(b)** and **Figure 5(b)** was charged with charges with a minus sign.

Recall that the signs of magnetic charges on the test plates were indicated in accordance with the poles of magnet during the separation of charges in pairs ±g (see **Figure 2**).

Along the vertical axis in **Figure 4** and **Figure 5**, the number of positive and negative counts ±n is deferred, which were determined in accordance with the generally accepted logic of electrostatic interactions. Measurements falling within the error zone (±10 mm) were considered as negative results. The reduction in the number of measurements in the case of studies of interactions in sets of three plates, is due to the fact that an excess on 5 positive metering over negative metering was achieved with their total number being reduction. In addition, in the case of set No. 1 (see the first set on the left), when the number of negative and positive reports turned out to be approximately equal, the total count of metering in the cycle was limited to 20 ± 2.

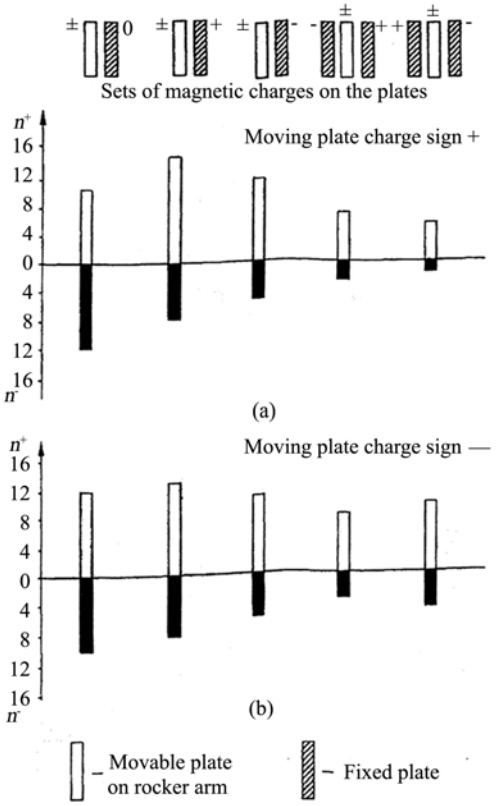


Figure 4. The effects magnetostatic interaction of copper plates charged with magnetic charges.

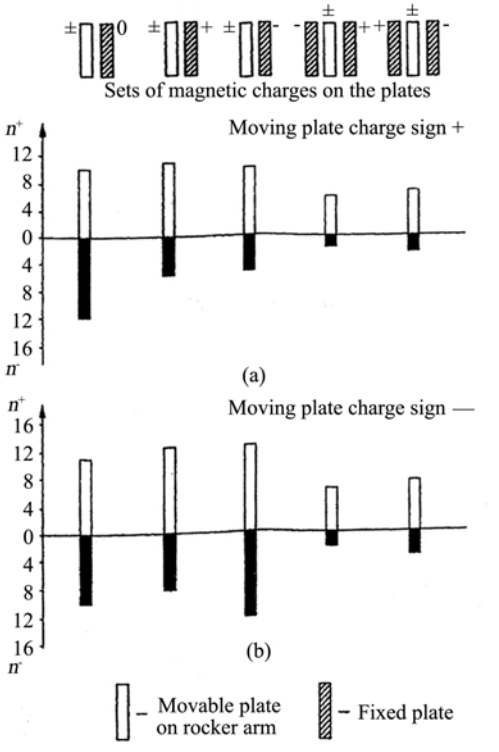


Figure 5. The effects magnetostatic interaction of nickel plates charged with magnetic charges.

Note 4. It is necessary to note an important circumstance that determined the general policy of this study. The fact is that the deviations of the rocker arm, when repeating the experiment within the same set of interacting plates, was differ significantly (up to 30%). As for the sign of interaction, *i.e.* direction of deflection of the charged plate on the rocker, then in this part everything was clearly reproduced. It was the latter circumstance that forced the author to limit the representative results only the sign of interaction. Possible causes of the noted anomalies are discussed below in Section 4.

The choice for study magnetic charges on nickel plates which is the ferromagnetic in their magnetic properties is associated with an attempt to determine the existence of the magnetic (diamagnetic) susceptibility in its composition. As noted above, it is magnetic dipoles which are free from inclusion in atomic structures and reside in potential conduction zones of solids that are responsible for the diamagnetic susceptibility of substance. As shown by the author's experiments with nickel the ferromagnets are no exception.

The irradiation by photons the plates charged of magnetic charges, with using light sources 12, had a weak effect on the interaction (a possible increase in the interaction forces in this case can be estimated by no more than ~10% - 15%) which significantly inferior to the effects observed in the Ehrenhaft experiments. The possible reason for this there is both a weakness of the light flux and a very limited access of photons to magnetic charges on the plates due to the geometry of our experiment.

In addition to copper and nickel plates, in the present experiments, on the subject of real magnetic charges, studied plates of lead and cadmium. The study of the magnetic charge of the last plates was carried out according to an abbreviated program, in which the interactions of only three charged plates with the magnetic charge of the plate on the on rocker of only one sign (minus) were evaluated. The results of the study of the magnetic charges on the lead and cadmium plates are presented in Figure 6.

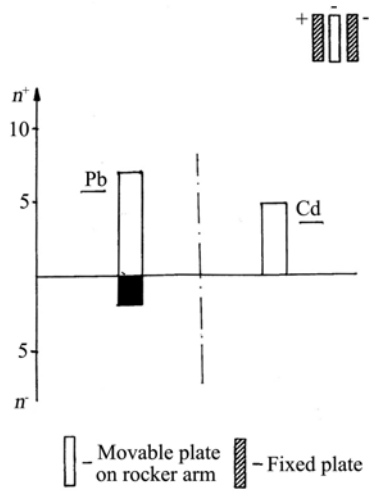


Figure 6. The effects magnetostatic interaction of magnetic-charged the lead and cadmium plates.

It is easy to see that the effects of the magnetostatic interaction between magnetic charges on lead and cadmium plates turned out to be close to what was observed under similar conditions, for example, for copper plates (see **Figure 4(b)**). At the same time, magnetostatic interactions between cadmium plates turned out to be the most intense, within the framework of which no negative readings were observed. It was in the case of cadmium plates in the experiment that the maximum deviation of the light spot on the scale-screen was noted, which amounted to 12 ± 1 cm.

4. The Discussion of the Results

The results of the experiments presented in the article, as well as the data of the author's subsequent research, are a serious evidence the existence of fundamental magnetic particles (real magnetic charges) in the composition of substance. It is magnetic charges, according to the conclusions of the author, that are the material filling of such a sacramental concept as the magnetic moment of an electron. In this case, it can be assumed that these charges may have the potential to realize some geometric autonomy within the framework of an atomic device. Under the influence of intra-crystalline fields that developed under the conditions of ionic composition in the lattice of the studied ferrites, physical conditions can be realized for the noted shifts in the magnetic charge density. Of course, the physics of this phenomenon is very exotic and will require special research to clarify it. Here it is important to keep in mind that the question concerns magnetic, not electric charges.

As for the author's position on the issue of the presented density shifts "MM", the main significance of this result, which is made on the basis of neutron-structural analysis using the trial and error method, he sees in the initiation of attention to the problem of the possible existence of real magnetic charges in the structures of substance. It is the results of the magnetic scattering of neutrons in ferrite structures and the author's conclusions about the "MM" shifts presented in the article that served as a kind of "compass" in planning, developing and conducting subsequent direct experimental searches for magnetic charges in a substance.

So, as part of a direct study of magnetic charges in matter presented in the article, charging of test bodies with magnetic charges of different signs was realized, what can be defined as the process of **magnetization** of test bodies. This process is a magnetic analogue of the process of **electrization** of a substance. However, the forces of magnetic charges confinement in matter, which is substantially strong than forces electron confinement, is the main reason for the small magnitude of magnetic charges on test bodies with which the author dealt in the described experiments [12]. It is the special conditions of the confinement of magnetic charges in substance are responsible the vicious ignoring by physical science the magnetic poles (magnetic charges) that reality exist in atoms and substance. In addition of magnetic charges, the real antielectrons, *i.e.* spinor par-

ticles with e+ charges, which were replaced in the theory by such surrogates as electronic vacancies or Dirac holes, are also in under harsh "sanctions" of confinement [13].

Thus, in fact, the entire existing physical theory is constructed using electric spinor particles with a negative charge (most often electrons). As for magnetic spinor particles, as well as electric particles with a positive charge, which make up threequarters of all spinor (charged) particles in the real World, all of them were replaced in the theory by surrogates noted by the author in this article, as well as discussed in detail in [14] [15].

In connection with the results of our experiments, we can assume that the magnetic charges on the particles of substance that research F. Ehrenhaft, were also the result of a process of mechanical magnetization, but in the Earth's magnetic field. The fact is that Ehrenhaft produced very small particles of substance for his experiments by mechanical grinding of solids, which was carried out in the Earth's magnetic field. Of course, the Earth's magnetic field is very small, but and the particles themselves and their magnetic charges were also extremely small. In his publications, Ehrenhaft emphasized that the search of particles with sufficient magnetic charge to conduct research, was a serious problem for him.

It is important to note here that the minimum magnetic charge on the particles, which was observed in the experiments of Ehrenhaft, was close to the charge of the electron. Thus, the formation of the magnetic charge on the test masses in the technology of the present author and F. Ehrenhaft, in principle, are analogous. The differences here are only in the scale of these processes and some experimental details.

The crucial parameter that was put by the author at the head of the angle of the present study is the sign or direction of the deviation of the charged plate from the zero position of the measuring system under the conditions of magnetostatic interaction. As for the magnitude of the forces that were realized in these experiments, they, based on an analysis of the elasticity of the tungsten filament used within the available angle of rotation of the rocker (5), could correspond to a range of $\sim 10^{-8}$ - 10^{-6} dyne. This is approximately 100 times more than the forces observed in the Ehrenhaft experiments which were estimated at $\sim 10^{-9}$ dyne and carried out in the atmosphere. The noted correlations of forces make it possible to calculate the possible number of elementary magnetic charges on particles in the Ehrenhaft experiments and on the test plates in our experiment. Of course, these are very rough estimates, but it turns out that in the Ehrenhaft experiments the number of elementary charges on the particles could be in the range from 1 до 10^2. As for the magnetic charges on the test plates, their number can be in the range 10^3 - 10^4. Recall that the magnitudes of the magnetic charges involved in the experiments, according to the author of the article, meet the condition e = g, where e is the electron charge.

It is also important to note that the author's technology for the separation of magnetic charges in pairs ±g allowed to charge test bodies with magnetic charges of any of the signs, by choice, which, apparently, is be done for the first time.

A well-known parameter that characterizes the diamagnetic properties of a substance is the magnetic susceptibility χ. It is the magnetic susceptibility, which, according to the author, can be called diamagnetic susceptibility, explains, for example, the effect of the excellent magnetostatic interaction observed for cadmium in relation, for example, to copper. The values of χ of copper are −0.086, cadmium −0.18, lead −1.1. In connection with the above-mentioned magnetostatic experiments with nickel plates, a very interesting result is the manifestation of the magnetic (diamagnetic) susceptibility of nickel, the magnitude of which is at the level of the magnetic susceptibility of copper and which very approximately can determined as 0.09 ± 0.03.

It should be noted one more interesting result that the present study has provided. As part of the set of plates No.1, *i.e.* when the interaction of a charged and uncharged plate was investigated, unlike other sets, no pronounced interaction effect was observed. This circumstance can be explained by the absence or very weak effect of the induction of magnetic charges in an uncharged plate under the influence of a magnetic charge of the adjacent plate. According to the author's concept, the magnetic charges in a normal conductor populate exclusively potential conduction zones and cannot go into real conduction zones.

In the Experimental Results section, significant differences were noted in the in deviations of the plate on the rocker when the experiment was repeated for the same set of plates. The reasons for this can be both anomalies in the mechanics of the tungsten filament during its deformation during the interaction, and a change in the degree of vacuum in the working chamber. The latter is quite possible, since the measurement time in each cycle was very significant, and the fore-vacuum pump was switched off during the measurement. In this case, the pressure in the chamber was maintained by a clamp, which clamped the vacuum hose.

The possibility of the influence of electric charges on the effects in the observed interactions is very small, given the careful grounding of all metal parts of the installation and test plates, in particular. So, for example, a charged copper plate showed practically the same magnetostatic effect after lying in a metal container for more than 2 weeks. The latter circumstance is explained by the fact that the magnetic charges with which the author worked in the above experiments populate, exclusively the potential conduction zones of solids. In this case, it is necessary to take into account the fact that the retention forces of magnetic charges in a substance are many times greater than the retention forces, for example, of electrons. As the author constantly emphasizes in his publications, the exit of magnetic charges into real conduction zones is realized only in superconductors.

Note 5. It is useful to note here that the implementation of the scenario described above with real magnetic charges in a substance suggests that the effect of superconductivity is determined precisely by magnetic charges that are unrecognized in physics, which at ultra-low temperatures transfer from potential in the real conduction zones of solids and, practically, are lose connection with the

lattice of superconductor. In this case, the electric charges in the electric current move in the conductor almost without resistance, because magnetic charges, with which particles of electric current must interact, do not experience "friction" against a lattice of superconductor.

5. Conclusions

This publication is an extended English-language presentation of the message presented in the author's Russian-language books, as well as in his English-language article (see [7]). The experiments presented in the article, as well as other experimental studies of the author, which are planned to be presented in his subsequent publications, are a serious basis for the assertion that magnetic charges are real structural components of atoms and substance.

Enormous support to the author in the development of his concept of real magnetic charges in matter was provided by the experiments and conclusions of F. Ehrenhaft and his followers. Given that only publications of the results of Ehrenhaft's research are dated from the period 1910-1951, it can be argued that this wonderful person put his whole life dedicated evidence of the existence the real magnetic charges in substance.

Based on the results of his experiments, as well as on the data of F. Ehrenhaft's experiments, the author developed the concept of the existence of real magnetic charges as structural components of atoms and substance, which he first presented to the scientific community in 2001 in a Russian-language book [5]. From 2001 to the present, the author has published 25 messages in Russian and English, in which he presented both his experiments and the concept of General physics with real magnetic charges in atoms and substance developed on their basis. In addition, the results of the author's research allowed him to make important amendments to the concepts of the real chemical bond physics, which significantly change many basic chemical concepts [16].

It should also be noted that the author's English-language articles, predominantly, are made in the so-called open journals, the content of which high physical theory tries to ignore. The author's attempts to publish his articles on real magnetic charges in high publishing houses thwarted by reviewers, and often simply by the editors of magazines, with the general wording: "this can never be".

The results of the author's research, presented in the article, as well as the results of F. Ehrenhaft's research, give an idea about serious experimental problems, associated with the study of real magnetic charges in the structures of substance. Namely, these problems were the reason that these fundamental particles, which make up half of all real fundamental charged particles in the real world, were not timely introduced into basic physics and are absent in the Standard Model.

However, it was impossible to completely isolate oneself from magnetic particles and true antielectrons in physical theory. Therefore, the various surrogates substitutes ("fig sheets") appeared, such as magnetic moments measured in the

magnetons of Bohr, 3d and 4f electrons, Dirac's holes and monopoles, electron shells of atoms and others. At the same time, it should be noted that world physical theory, relying solely on electric particles and while completely ignoring magnetic particles, accomplished the impossible, *i.e.* was able to more or less adequately explain the existing physical reality. As a result, theoretical "Himalayas" were built from all kinds of super complex physical concepts which under the conditions of the existence of real magnetic charges and true antielectrons in atoms and substance should be radically revised.

The author hopes that the datura, in the form of electric magnetism [17] [18], which prevents the recognition of real magnetic poles and, for almost 150 years, decomposes a physical theory, will disappear and magnetic charges which are real components of atoms and substance will receive long-awaited "citizenship" on planet Earth.

Acknowledgements

The author expresses deep gratitude to **Elena Sizova** and **Vitaly Sizov** who for many years helped and supported him which largely contributed to the implementation of his research activities.

The author is grateful to the technical specialists: **Alexander Davydov, Nikolai Koshelev, Alexander Lomakin, Evgeni Moiseenkov and Sergei Staritsyn** who together with him conducted the main experiments connected with the discovery and investigation of real magnetic charges.

The author is grateful to Professor **V. I. Panov** (Moscow State University), who back in 1975 advised the author on the design and methods of working with highly sensitive torsion scales. It is on the advice of V. I. Panov in the contruction of torsion scales, the author used a tungsten filament.

Conflicts of Interest

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

References

[1] Erenhaft, F. (1910) Wiener Berichte 119(11a) 836;
(1942) Jour. Frank. Inst. Mar., 235;
(1951) Acta Phys. Austriaca, 5, 12.

[2] Mikhailov, V.F. and Ruzicka, J. (1989) *Acta Physica Universitatis Comenianae*, **29**, 97-148.

[3] Sizov, R.A. (1971) *Journal of Experimental and Theoretical Physics*, **60**, 1363-1370.

[4] Sizov, R.A. (1970) *Solid State Physics*, **13**, 2081-2088.

[5] Sizov, R.A. (2001) New Presentation of Nature Magnetism, Gravitation and Nuclear Forces of Bonding. Akademizdat Center "Science", Moscow, 62 p.

[6] Sizov, R.A. (2008) Electric and Magnetic Spinor Particles as Structure Forming Components of Mass and Electromagnetic Source Gravitation. Akademizdat Center "Science", Moscow, 260 p.

[7] Sizov, R.A. (2015) *Journal of Modern Physics*, **6**, 1013-1022. https://doi.org/10.4236/jmp.2015.68106

[8] Sizov, R.A. (1971) *Soviet Physics JETF*, **33**, 737-740.

[9] Sizov, R.A. (2016) *Journal of Modern Physics*, **7**, 2374-2397. https://doi.org/10.4236/jmp.2016.716206

[10] Dirac, P.A.M. (1931) *Proceedings of the Royal Society A*, **133**, 60-72. https://doi.org/10.1098/rspa.1931.0130

[11] Sizov, R.A. (2018) *Physics & Astronomy International Journal*, **2**, Article No. 00043.

[12] Sizov, R.A. (2017) *Journal of Modern Physics*, **8**, 1072-1086. https://doi.org/10.4236/jmp.2017.87069

[13] Dirac, P.A.M. (1930) *Proceedings of the Royal Society A*, **126**, 360. https://doi.org/10.1098/rspa.1930.0013

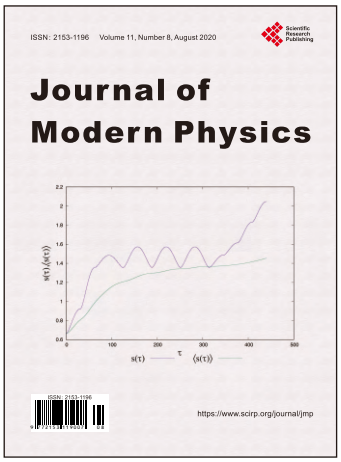
[14] Sizov, R.A. (2015) *Journal of Modern Physics*, **6**, 2280-2289. https://doi.org/10.4236/jmp.2015.615232

[15] Sizov, R.A. (2019) *Applied Science and Innovative Research*, **3**, 257-263. https://doi.org/10.22158/asir.v3n4p257

[16] Sizov, R.A. (2020) *Open Journal of Physical Chemistry*, **10**, 48-67. https://doi.org/10.4236/ojpc.2020.101003

[17] Sizov, R.A. (2019) *Applied Physics Research*, **11**, 49-55. https://doi.org/10.5539/apr.v11n3p49

[18] Sizov, R.A. (2020) *OSP Journal of Nuclear Science*, **2**, JNS-2-111.



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