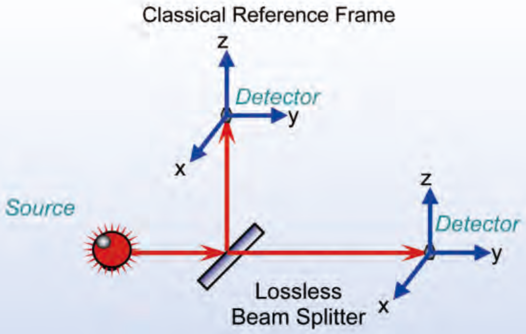
ISSN: 2153-1196 Volume 9, Number 14, December 2018

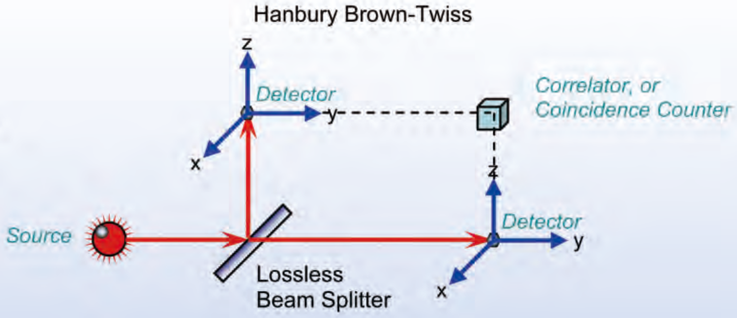
Scientific
Research
Publishing

Journal of Modern Physics



Classical Reference Frame

and



Hanbury Brown-Twiss

ISSN: 2153-1196

www.scirp.org/journal/jmp

Journal Editorial Board

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

Editor-in-Chief

Prof. Yang-Hui He	City University, UK

Editorial Board

Prof. Nikolai A. Sobolev	Universidade de Aveiro, Portugal
Dr. Mohamed Abu-Shady	Menoufia University, Egypt
Dr. Hamid Alemohammad	Advanced Test and Automation Inc., Canada
Prof. Emad K. Al-Shakarchi	Al-Nahrain University, Iraq
Prof. Tsao Chang	Fudan University, China
Prof. Changle Chen	University of Science and Technology of China, China
Prof. Stephen Robert Cotanch	NC State University, USA
Prof. Peter Chin Wan Fung	University of Hong Kong, China
Prof. Ju Gao	The University of Hong Kong, China
Prof. Sachin Goyal	University of California, USA
Dr. Wei Guo	Florida State University, USA
Prof. Cosmin Ilie	Los Alamos National Laboratory, USA
Prof. Haikel Jelassi	National Center for Nuclear Science and Technology, Tunisia
Prof. Santosh Kumar Karn	Dr. APJ Abdul Kalam Technical University, India
Prof. Christophe J. Muller	University of Provence, France
Prof. Ambarish Nag	National Renewable Energy Laboratory, USA
Dr. Rada Novakovic	National Research Council, Italy
Prof. Tongfei Qi	University of Kentucky, USA
Prof. Mohammad Mehdi Rashidi	University of Birmingham, UK
Prof. Alejandro Crespo Sosa	Universidad Nacional Autónoma de México, Mexico
Dr. A. L. Roy Vellaisamy	City University of Hong Kong, China
Prof. Yuan Wang	University of California, Berkeley, USA
Prof. Fan Yang	Fermi National Accelerator Laboratory, USA
Prof. Peter H. Yoon	University of Maryland, USA
Prof. Meishan Zhao	University of Chicago, USA
Prof. Pavel Zhuravlev	University of Maryland at College Park, USA

Journal of Modern Physics, 2018, 9, 2355-2724
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Table of Contents

Volume 9 Number 14 December 2018

Journal of Modern Physics, 2018, 9, 2355-2724
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

The figure on the front cover is from the article published in Journal of Modern Physics, 2018, Vol. 9, No. 14, pp. 2378-2390 by Daniel Sepunaru.

Journal of Modern Physics (JMP)
Journal Information

SUBSCRIPTIONS

The *Journal of Modern Physics* (Online at Scientific Research Publishing, 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, 2018, 9, 2355-2369
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Hamiltonian, Path Integral and BRST Formulations of the Restricted Gauge Theory of *QCD_2*

Usha Kulshreshtha[1*], Daya Shankar Kulshreshtha[2#], James P. Vary[3]

[1]Department of Physics, Kirori Mal College, University of Delhi, Delhi, India
[2]Department of Physics and Astrophysics, University of Delhi, Delhi, India
[3]Department of Physics and Astronomy, Iowa State University, Ames, Iowa, USA
Email: *ushakulsh@gmail.com, dskulsh@gmail.com, jvary@iastate.edu

How to cite this paper: Kulshreshtha, U., Kulshreshtha, D.S. and Vary, J.P. (2018) Hamiltonian, Path Integral and BRST Formulations of the Restricted Gauge Theory of QCD_2. *Journal of Modern Physics*, **9**, 2355-2369.
https://doi.org/10.4236/jmp.2018.914150

Received: September 13, 2018
Accepted: December 2, 2018
Published: December 5, 2018

Abstract

We study the Hamiltonian, path integral and Becchi-Rouet-Stora and Tyutin (BRST) formulations of the restricted gauge theory of QCD_2 à la Cho *et al.* under appropriate gauge-fixing conditions.

Keywords

Hamiltonian Quantization, Path Integral Quantization, BRST Quantization, Quantum Chromodynamics, QCD_2, Field Theories in Lower Dimensions, Gauge-Invariant Theories, Gauge-Fixing

1. Introduction

In this work we consider the restricted gauge theory of quantum chromodynamics (QCD) in one-space one-time dimension (QCD_2) à la Cho *et al.* [1]-[14], studied rather widely [2]-[23], and study its quantization using Hamiltonian [24], path integral [25] [26] [27] [28] and Becchi-Rouet-Stora and Tyutin (BRST) [29] [30] [31], formulations [24]-[31], in the usual instant-form (IF) of dynamics (on the hyperplanes: $x^0 = t = \text{constant}$) [32] [33]. We recap the basis of this theory in the next section where we also highlight the motivations for the present study. The theory is seen to be gauge-invariant (GI) possessing a set of first-class constraints [14]. We quantize this theory under appropriate gauge-fixing conditions (GFC's) using the Hamiltonian and path integral formulations [24] [25] [26] [27]

[#]Part of this work was presented by DSK as Invited Contributed Talk at the International Conference on Light-Cone Physics: LC2014, held at North Carolina State University, Rayleigh, NC, USA, during May 26-30, 2014.

[28].

However, in the usual Hamiltonian and path integral quantization [24] [25] [26] [27] [28] of a theory under some GFC's, the gauge-invariance of the theory gets broken because of the gauge-fixing. In view of this, in order to achieve the quantization of the theory such that the gauge-invariance of the theory is maintained even under gauge-fixing, we go to a more generalized procedure called the BRST quantization [29] [30] [31], [27] [28] [29] [30] [31], where the extended gauge symmetry of the theory (called the BRST symmetry) is maintained even under gauge-fixing. In fact, this also necessitates a study of the BRST quantization of the theory to achieve a kind of complete quantization of the theory.

The paper is organized as follows. In the next section, we briefly recap the basics of the so-called restricted gauge theory of QCD_2 à la Cho *et al.* [1]-[13]. In Section 3, we study its Hamiltonian and path integral formulations. Its BRST formulation is studied in Section 4. Finally the summary and discussion is given in Section 5.

2. Restricted Gauge Theory of QCD_2 à la Cho *et al.*: A Recap of Basics

In this section, we recap the basics of the restricted gauge theory of QCD_2 à la Cho *et al.* [1]-[14] and others [1]-[23]. The theory makes use of the so-called "Cho-decomposition", which is, in fact, the gauge independent decomposition of the non-Abelian potential into the restricted potential and the valence potential and it helps in the clarification of the topological structure of the non-Abelian gauge theory, and it also takes care of the topological characters in the dynamics.

The non-Abelian gauge theory has rich topological structures manifested by the non-Abelian monopoles, the multiple vacua and the instantons and one needs to take into account these topological characters in the non-Abelian dynamics. Since the decomposition of the non-Abelian connection contains these topological degrees explicitly, it can naturally take care of them in the non-Abelian dynamics.

An important consequence of the decomposition is that it allows one to view QCD as the restricted gauge theory (made of the restricted potential) which is coupled to a gauge-covariant colored vector field (the valence potential). The restricted potential is defined in such a way that it allows a covariantly constant unit isovector $\hat{n}$ everywhere in space-time, which enables one to define the gauge-independent color direction everywhere in space-time and, at the same time, allows one to define the magnetic potential of the non-Abelian monopoles. Furthermore it has the full SU(2) gauge degrees of freedom, in spite of the fact that it is restricted. Consequently, the restricted QCD made of the restricted potential describes a very interesting dual dynamics of its own, and plays a crucial role in the understanding of QCD.

On the other hand, the restricted QCD is a constrained system, due to the presence of the topological field $\hat{n}$ which is constrained to have the unit norm. A natural way to accommodate the topological degrees into the theory is to introduce a topological field $\hat{n}$ of unit norm, and to decompose the connection into the Abelian projection part which leaves $\hat{n}$ a covariant constant and the remaining part which forms a covariant vector field:

$$\begin{aligned} &\mathbf{A}_\mu = A_\mu \hat{n} - \frac{1}{g}\hat{n}\times\partial_\mu \hat{n} + \mathbf{X}_\mu = \hat{A}_\mu + \mathbf{X}_\mu \\ &A_\mu = \hat{n}\cdot\mathbf{A}_\mu \\ &\hat{n}^2 = 1 \end{aligned} \tag{1a}$$

where A^μ is the "electric" potential and the Abelian projection A^μ is precisely the connection which leaves $\hat{n}$ invariant under the parallel transport and makes $\hat{n}$ a covariant constant:

$$\hat{D}_\mu \hat{n} = \partial_\mu \hat{n} + g\hat{A}_\mu \times \hat{n} = 0 \tag{2}$$

Also, under the infinitesimal gauge-transformation:

$$\begin{aligned} &\delta\mathbf{A}_\mu = \frac{1}{g}D_\mu\vec{\alpha}\ ,\ \delta A_\mu = \frac{1}{g}\hat{n}\cdot\partial_\mu\vec{\alpha}\ ,\ \delta\hat{A}_\mu = \frac{1}{g}\hat{D}_\mu\vec{\alpha} \\ &\delta\mathbf{X}_\mu = -\vec{\alpha}\times\mathbf{X}_\mu\ ,\ \delta\hat{n} = -\vec{\alpha}\times\hat{n} \end{aligned} \tag{3}$$

This shows that A^μ by itself describes an $SU(2)$ connection which enjoys the full $SU(2)$ gauge degrees of freedom. The restricted potential A^μ is defined by the Abelian projection and the connection space (the space of all gauge potentials) forms an affine space. Indeed the affine nature of the connection space guarantees that one can describe an arbitrary potential simply by adding a gauge-covariant piece X_μ to the restricted potential.

The above mentioned decomposition is known as the Cho-decomposition or the Cho-Faddeev-Niemi decomposition. It was introduced [2]-[23] in an attempt to demonstrate the monopole condensation in QCD. The decomposition itself and the importance of this decomposition in clarifying the non-Abelian dynamics in QCD_2 has been studied by many authors and for further details we refer to the work of Refs. [2]-[23].

The restricted potential A^μ actually has a dual structure and the field strength made of the restricted potential is decomposed as:

$$\hat{F}_{\mu\nu} = \left(F_{\mu\nu} + H_{\mu\nu}\right)\hat{n},\ F_{\mu\nu} = \left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) \tag{4a}$$

$$H_{\mu\nu} = -\frac{1}{g}\hat{n}\cdot\left(\partial_\mu\hat{n}\times\partial_\nu\hat{n}\right) = \left(\partial_\mu\tilde{C}_\nu - \partial_\nu\tilde{C}_\mu\right) \tag{4b}$$

where C_μ is the "magnetic" potential. Further, following the work of Refs. [2]-[23], it is possible to introduce the magnetic potential as above (at least locally section-wise) in view of the following identity:

$$\partial_\mu\tilde{H}_{\mu\nu} = 0,\ \tilde{H}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}H_{\rho\sigma} \tag{5}$$

which allows one to identify the non-Abelian monopole potential by:

$$\mathbf{C}_\mu = -\frac{1}{g}\hat{n}\times\partial_\mu\hat{n} \tag{6}$$

in terms of which the magnetic field is expressed as

$$\mathbf{H}_{\mu\nu} = \partial_\mu\mathbf{C}_\nu - \partial_\nu\mathbf{C}_\mu + g\mathbf{C}_\mu\times\mathbf{C}_\nu = -\frac{1}{g}\partial_\mu\hat{n}\times\partial_\nu\hat{n} = H_{\mu\nu}\hat{n} \tag{7}$$

With the above connection (-albeit decomposition) one has:

$$\begin{aligned}&\mathbf{F}_{\mu\nu} = \left[\left(F_{\mu\nu}+H_{\mu\nu}\right)\hat{n} + \hat{D}_\mu\mathbf{X}_\nu - \hat{D}_\nu\mathbf{X}_\mu + g\mathbf{X}_\mu\times\mathbf{X}_\nu\right]\\ &\hat{n}\cdot\mathbf{X}_\mu = 0\\ &\hat{n}\cdot\hat{D}_\mu\mathbf{X}_\nu = 0\end{aligned} \tag{8}$$

and for the Yang-Mills Lagrangian density one has: [1]-[15]:

$$\begin{aligned}\tilde{\mathcal{L}} = \Big[&-\frac{1}{4}\hat{F}_{\mu\nu}\hat{F}^{\mu\nu} - \frac{1}{4}\left(\hat{D}_\mu\mathbf{X}_\nu - \hat{D}_\nu\mathbf{X}_\mu\right)^2 - \frac{g}{2}\hat{F}_{\mu\nu}\cdot\left(\mathbf{X}_\mu\times\mathbf{X}_\nu\right)\\ &-\frac{g^2}{4}\left(\mathbf{X}_\mu\times\mathbf{X}_\nu\right)^2 + \lambda\left(\hat{n}^2-1\right) + \lambda_\mu\hat{n}\cdot\mathbf{X}_\mu\Big]\end{aligned} \tag{9}$$

where λ and λ_μ are the Lagrange multiplier fields and $g^{\mu\nu} := g_{\mu\nu} = diag(+1,-1)$. The Lagrangian density of the so-called restricted gauge theory made of the Abelian projection without $\mathbf{X}_\mu$ is therefore defined by [2]-[14]:

$$\mathcal{L} = \left[-\frac{1}{4}\hat{F}_{\mu\nu}\hat{F}^{\mu\nu} + \lambda\left(\hat{n}^2-1\right)\right] \tag{10}$$

The theory defined by the above Lagrangian density has a full $SU(2)$ gauge invariance and it describes the dual dynamics of QCD with the dynamical degrees of the maximal Abelian subgroup U(1) as the electric component and the topological degrees of SU(2) as the magnetic component. It therefore represents an important model in the QCD theory namely, in QCD_2 and deserves to be studied more properly. One of the important steps in this direction is to construct the quantum theory corresponding to this classical theory of QCD_2 by quantizing the theory. This provides motivation for the present studies and in fact, necessitates our presents studies. In the next section, we consider the Hamiltonian and path integral formulations of this constrained theory.

3. Hamiltonian and Path Integral Formulations

We now study the Hamiltonian and path integral quantization of the above restricted gauge theory of QCD_2 (made of the Abelian projection without $\mathbf{X}_\mu$) defined by the Lagrangian density [2]:

$$\begin{aligned}\mathcal{L} &= \left[-\frac{1}{4}\hat{F}_{\mu\nu}\hat{F}^{\mu\nu} + \lambda\left(\hat{n}^2-1\right)\right] = \left[-\frac{1}{4}G_{\mu\nu}G^{\mu\nu} + \lambda\left(\hat{n}^2-1\right)\right]\\ &= \left[-\frac{1}{4}\left(F_{\mu\nu}+H_{\mu\nu}\right)\left(F^{\mu\nu}+H^{\mu\nu}\right) + \lambda\left(\hat{n}^2-1\right)\right]\\ &= \left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}F_{\mu\nu}H^{\mu\nu} - \frac{1}{4}H_{\mu\nu}H^{\mu\nu} + \lambda\left(\hat{n}^2-1\right)\right]\end{aligned} \tag{11}$$

In the instant-form (IF) of dynamics, the above Lagrangian density reads:

$$\mathcal{L}=\left[\frac{1}{2}\left(\partial_0 A_1-\partial_1 A_0\right)^2+\frac{1}{g}\left(\partial_0 A_1-\partial_1 A_0\right)\Phi+\frac{1}{2g^2}\Phi^2+\lambda\left(\hat{n}^2-1\right)\right] \tag{12a}$$

$$\Phi=\left[\partial_0\hat{n}\cdot\left(\hat{n}\times\partial_1\hat{n}\right)\right] \tag{12b}$$

Here Φ is another topological scalar field constructed out of the topological field $\hat{n}$ (of unit norm) and its space derivative ($\partial_1\hat{n}$) as well as its time derivative ($\partial_0\hat{n}$), as defined above. The canonical momenta obtained for the above theory are:

$$\Pi_\lambda=\frac{\partial\mathcal{L}}{\partial\left(\partial_0\lambda\right)}=0,\ \Pi^0=\frac{\partial\mathcal{L}}{\partial\left(\partial_0 A_0\right)}=0 \tag{13a}$$

$$E\left(:=\Pi^1\right)=\frac{\partial\mathcal{L}}{\partial\left(\partial_0 A_1\right)}=\left(\partial_0 A_1-\partial_1 A_0\right)+\frac{1}{g}\Phi \tag{13b}$$

$$\hat{\Pi}_{\hat{n}}=\frac{\partial\mathcal{L}}{\partial\left(\partial_0\hat{n}\right)}=-\frac{1}{g}\left(\partial_0 A_1-\partial_1 A_0\right)\left(\hat{n}\times\partial_1\hat{n}\right) \tag{13c}$$

Here Π_λ, Π^0, $E\left(:=\Pi^1\right)$, and $\hat{\Pi}_{\hat{n}}$ are the momenta conjugate canonically to λ, A_0, A_1, and $\hat{n}$ respectively. The above equations however, imply that the theory possesses three primary constraints:

$$\chi_1=\Pi_\lambda\approx 0$$

$$\Omega_1=\Pi^0\approx 0$$

$$\psi_1=\hat{n}\cdot\hat{\Pi}_{\hat{n}}\approx 0 \tag{14}$$

The symbol $\approx$ here denotes a weak equality in the sense of Dirac [24]. The canonical Hamiltonian density of the theory $\mathcal{H}_c$ is:

$$\mathcal{H}_c^N=\left[\frac{1}{2}E^2+E\partial_1 A_0-\lambda\left(\hat{n}^2-1\right)\right] \tag{15}$$

After including the primary constraint χ_1, Ω_1, and ψ_1 in the canonical Hamiltonian density $\mathcal{H}_c$ with the help of Lagrange multiplier field $u(x,t)$, $v(x,t)$ and $w(x,t)$ which is treated as dynamical, the total Hamiltonian density of the theory $\mathcal{H}_T$ could be written as:

$$\mathcal{H}_T^N=\left[\frac{1}{2}E^2+E\partial_1 A_0-\lambda\left(\hat{n}^2-1\right)+\chi_1 u+\Omega_1 v+\psi_1 w\right] \tag{16}$$

The Hamilton's equations of motion of the theory that preserve the constraints of the theory in the course of time could be obtained from the total Hamiltonian: $H_T=\int\mathcal{H}_T\mathrm{d}x^1$ and are omitted here for the sake of brevity. Demanding the preservation of the primary constraints χ_1 and Ω_1 for all time leads to the secondary Gauss-law constraints χ_2 and Ω_2 respectively:

$$\chi_2=\left(\hat{n}^2-1\right)\approx 0$$

$$\Omega_2=\partial_1 E\approx 0 \tag{17}$$

The preservation of ψ_1 for all times does not lead to any secondary constraint. The preservation of χ_2 and Ω_2 also does not lead to any new constraints. The theory is therefore seen to possess a set of five constraints:

$$\begin{aligned}\rho_1 &= \Omega_1 = \Pi^0 \approx 0\\ \rho_2 &= \Omega_2 = \partial_1 E \approx 0\\ \rho_3 &= \chi_1 = \Pi_\lambda \approx 0\\ \rho_4 &= \chi_2 = \left(\hat{n}^2 - 1\right) \approx 0\\ \rho_5 &= \psi_1 = \hat{n}\cdot\hat{\Pi}_{\hat{n}} \approx 0\end{aligned} \tag{18}$$

Matrix $R_{\alpha\beta}$ of the Poisson brackets of the above constraints ρ_i among themselves is clearly singular implying that the set of these constraints ρ_i is first-class and that the theory under consideration is GI. In fact, the theory is seen to be invariant under the local vector gauge transformations:

$$\begin{gathered}\delta A_\mu = \frac{1}{g}\hat{n}\cdot\partial_\mu\vec{\alpha} = \frac{1}{g}\partial_\mu\alpha\,,\, \delta\lambda = \frac{1}{g}\partial_0\alpha\,,\, \delta\hat{n} = \hat{n}\times\vec{\alpha} = 0\\ \delta\Pi^0 = \delta E = \delta\Pi_\lambda = \delta\hat{\Pi}_{\hat{n}} = \delta\Pi_u = \delta\Pi_v = \delta\Pi_w = 0\end{gathered} \tag{19}$$

where gauge parameter $\alpha(x,t)$ is an arbitrary function of its arguments. The components of the vector gauge current $\left(J^\mu \equiv J^0, J^1\right)$ are:

$$\begin{aligned}J^0 &= \int \mathrm{d}x^0\mathrm{d}x^1 j^0 = \int \mathrm{d}x^0\mathrm{d}x^1\left[\frac{1}{g}\partial_1\alpha\left(\partial_0 A_1 - \partial_1 A_0 + \frac{1}{g}\Phi\right)\right]\\ J^1 &= \int \mathrm{d}x^0\mathrm{d}x^1 j^1 = \int \mathrm{d}x^0\mathrm{d}x^1\left[-\frac{1}{g}\partial_0\alpha\left(\partial_0 A_1 - \partial_1 A_0 + \frac{1}{g}\Phi\right)\right]\end{aligned} \tag{20}$$

The theory is clearly gauge-invariant and could now be quantized under appropriate gauge-fixing conditions (GFC's), which could e.g. be chosen as (which by no means is an unique choice):

$$\zeta_1 = \lambda \approx 0,\ \zeta_2 = A_1 \approx 0,\ \zeta_3 = A_0 \approx 0 \tag{21}$$

It may be important to mention here that any set of GFC's could be chosen here such that the resulting set of constraints of the theory (including the set of GFC's) becomes a set of second-class constraints so that the matrix of the total set of constraints becomes non-singular and consequently could be inverted.

The total set of constraints of the theory under these GFC's then becomes:

$$\xi_1 = \rho_1 = \Omega_1 = \Pi^0 \approx 0 \tag{22a}$$

$$\xi_2 = \rho_2 = \Omega_2 = \partial_1 E \approx 0 \tag{22b}$$

$$\xi_3 = \rho_3 = \chi_1 = \Pi_\lambda \approx 0 \tag{22c}$$

$$\xi_4 = \rho_4 = \chi_2 = \left(\hat{n}^2 - 1\right) \approx 0 \tag{22d}$$

$$\xi_5 = \rho_5 = \psi_1 = \hat{n}\cdot\hat{\Pi}_{\hat{n}} \approx 0 \tag{22e}$$

$$\xi_6 = \zeta_1 = \lambda \approx 0 \tag{22f}$$

$$\xi_7 = \zeta_2 = A_1 \approx 0 \tag{22g}$$

$$\xi_8 = \zeta_3 = A_0 \approx 0 \tag{22h}$$

The non-vanishing matrix elements of the matrix $M_{\alpha\beta}$ of the Poisson Bracket's's among these above constraints ξ_i are:

$$M_{18} = -M_{81} = -\delta\left(x-y\right) \tag{23a}$$

$$M_{27} = +M_{72} = -\partial_1\delta\left(x-y\right) \tag{23b}$$

$$M_{36} = -M_{63} = -\delta\left(x-y\right) \tag{23c}$$

$$M_{45} = -M_{54} = +2\hat{n}^2\delta\left(x-y\right) \tag{23d}$$

The above matrix $M_{\alpha\beta}$ is clearly non-singular implying that the constraints ξ_i form a set of second-class constraints. The theory could therefore be quantized using the Dirac's Hamiltonian Formulation as well as using the path integral formulation. The square root of the absolute value of the determinant of this matrix $M_{\alpha\beta}$ is:

$$\left[\left\|det\left(M_{\alpha\beta}\right)\right\|\right]^{\frac{1}{2}} = \left[2\hat{n}^2\left(\delta\left(x-y\right)\right)^3\partial_1\delta\left(x-y\right)\right] \tag{24}$$

The non-vanishing equal-time Dirac brackets (DB's) of the theory are [24]:

$$\left\{\Pi^a_{\hat{n}}\left(x,t\right),\Pi^b_{\hat{n}}\left(y,t\right)\right\}_{DB} = \frac{1}{\hat{n}^2}\left(n^a\Pi^b_{\hat{n}} - \Pi^a_{\hat{n}}n^b\right)\delta\left(x-y\right) \tag{25a}$$

$$\left\{n^a\left(x,t\right),\Pi^b_{\hat{n}}\left(y,t\right)\right\}_{DB} = \frac{1}{\hat{n}^2}\left(n^a\Pi^b_{\hat{n}} - \Pi^a_{\hat{n}}n^b\right)\delta\left(x-y\right) \tag{25b}$$

$$\left\{\Pi^a_{\hat{n}}\left(x,t\right),n^b\left(y,t\right)\right\}_{DB} = \frac{-1}{\hat{n}^2}\left(n^a\Pi^b_{\hat{n}} - \Pi^a_{\hat{n}}n^b\right)\delta\left(x-y\right) \tag{25c}$$

Here one needs to remember that while making a transition from equal-time Dirac brackets to the equal-time commutation relations using the Dirac quantization rule, one needs to take in to account the problem of operator ordering (which occurs here because the results of the equal-time commutation relations involve the product of the operators). Also, the roman indices *a* and *b* here, are the color indices of the gauge theory of QCD_2.

Also for the later use, for considering the BRST formulation of the theory we convert the total Hamiltonian density of the theory into the first order Lagrangian density $\mathcal{L}_{IO}$:

$$\begin{aligned}\mathcal{L}_{IO} &= \Big[\Pi_\lambda\partial_0\lambda + \Pi^0\partial_0A_0 + E\partial_0A_1 + \hat{\Pi}_{\hat{n}}\cdot\partial_0\hat{n} \\ &\quad + \Pi_u\partial_0u + \Pi_v\partial_0v + \Pi_w\partial_0w - \mathcal{H}_T\Big] \\ &= \left[\frac{1}{2}\left(\partial_0A_1 - \partial_1A_0\right)^2 - \frac{1}{2g^2}\Phi^2 + \lambda\left(\hat{n}^2 - 1\right)\right]\end{aligned} \tag{26}$$

For considering the path integral formulation, the transition to quantum theory is made again by writing the vacuum to vacuum transition amplitude for the theory, called the generating functional $Z\left[J_k\right]$ of the theory, following again the Senjanovic procedure for a theory possessing a set of second-class constraints [25] [26] [27] [28], appropriate for our present theory, considered under the gauge-fixing conditions ξ_i, in the presence of the external sources: J_k as follows [25] [26] [27] [28]:

$$
\begin{aligned}
Z\left[J_k\right]=\int[\mathrm{d}\mu]\exp\Big[i\int\mathrm{d}x^0\mathrm{d}x^1\Big[&J_k\Phi^k+\Pi_\lambda\partial_0\lambda+\Pi^0\partial_0A_0+E\partial_0A_1\\
&+\hat{\Pi}_{\hat{n}}\cdot\partial_0\hat{n}+\Pi_u\partial_0u+\Pi_v\partial_0v+\Pi_w\partial_0w-\mathcal{H}_T\Big]\Big]
\end{aligned}
\tag{27}
$$

where the phase space variables of the theory are: $\Phi^k\equiv\left(\lambda,A_0,A_1,\hat{n},u,v,w\right)$ with the corresponding respective canonical conjugate momenta: $\Pi_k\equiv\left(\Pi_\lambda,\Pi_0,E,\hat{\Pi}_{\hat{n}},\Pi_u,\Pi_v,\Pi_w\right)$. The functional measure $[\mathrm{d}\mu]$ of the generating functional $Z\left[J_k\right]$ under this gauge-fixing is obtained as:

$$
\begin{aligned}
[\mathrm{d}\mu]=\Big[\Big[&2\hat{n}^2\left(\delta(x-y)\right)^3\partial_1\delta(x-y)\Big][\mathrm{d}\lambda][\mathrm{d}A_0][\mathrm{d}A_1][\mathrm{d}\hat{n}][\mathrm{d}u][\mathrm{d}v][\mathrm{d}w][\mathrm{d}\Pi_\lambda]\\
&\left[\mathrm{d}\Pi^0\right][\mathrm{d}E]\left[\mathrm{d}\hat{\Pi}_{\hat{n}}\right][\mathrm{d}p_u][\mathrm{d}p_v][\mathrm{d}p_w]\delta\left[\Pi^0\approx0\right]\delta\left[\partial_1E\approx0\right]\delta\left[\Pi_\lambda\approx0\right]\\
&\delta\left[\left(\hat{n}^2-1\right)\approx0\right]\delta\left[\hat{n}\cdot\hat{\Pi}_{\hat{n}}\approx0\right]\delta[\lambda\approx0]\delta[A_1\approx0]\delta[A_0\approx0]\Big]
\end{aligned}
\tag{28}
$$

This completes the Hamiltonian and path integral formulations of the theory. The BRST formulation of the theory is considered in the next section.

4. BRST Formulation

In the following, we study the BRST formulation of the theory. For the BRST formulation of the model, we rewrite the theory as a quantum system that possesses the generalized gauge invariance called BRST symmetry. For this, we first enlarge the Hilbert space of our gauge-invariant theory and replace the notion of gauge-transformation, which shifts operators by c-number functions, by a BRST transformation, which mixes operators with Bose and Fermi statistics. We then introduce new anti-commuting variables c and $\bar{c}$ (Grassman numbers on the classical level and operators in the quantized theory) and a commuting variable b such that [27] [28] [29] [30] [31]:

$$\hat{\delta}\lambda=\frac{1}{g}\partial_0c,\ \hat{\delta}A_0=\frac{1}{g}\partial_0c,\ \hat{\delta}A_1=\frac{1}{g}\partial_1c,\ \hat{\delta}u=\frac{1}{g}\partial_0\partial_0c \tag{29a}$$

$$\hat{\delta}\hat{\Pi}_{\hat{n}}=\hat{\delta}\Pi^0=\hat{\delta}E=\hat{\delta}\Pi_\lambda=\hat{\delta}\Pi_u=\hat{\delta}\Pi_v=\hat{\delta}\Pi_w=0 \tag{29b}$$

$$\hat{\delta}\hat{n}=\hat{\delta}v=\hat{\delta}w=0,\ \hat{\delta}c=0,\ \hat{\delta}\bar{c}=b,\ \hat{\delta}b=0 \tag{29c}$$

with the property $\hat{\delta}^2=0$. We now define a BRST-invariant function of the dynamical phase space variables of the theory to be a function f such that $\hat{\delta}f=0$. Now the BRST gauge-fixed quantum Lagrangian density $\mathcal{L}_{BRST}$ for the theory could be obtained by adding to the first-order Lagrangian density $\mathcal{L}_{IO}$, a trivial BRST-invariant function (e.g.) as follows:

$$
\begin{aligned}
\mathcal{L}_{BRST}=\Big[&\frac{1}{2}\left(\partial_0A_1-\partial_1A_0\right)^2-\frac{1}{2g^2}\Phi^2+\lambda\left(\hat{n}^2-1\right)\\
&+\hat{\delta}\left[\bar{c}\left(g\partial_0A_0+g\partial_1A_0-g\partial_0A_1+\frac{1}{2}b\right)\right]\Big]
\end{aligned}
\tag{30}
$$

The last term in the above equation is the extra BRST-invariant gauge-fixing term. After one integration by parts, the above equation could now be written as:

$$\begin{aligned}\mathcal{L}_{BRST} := \Bigg[& \frac{1}{2}\left(\partial_0 A_1 - \partial_1 A_0\right)^2 - \frac{1}{2g^2}\Phi^2 + \lambda\left(\hat{n}^2 - 1\right) \\ & + \frac{1}{2}b^2 + gb\left(\partial_0 A_0 - E + \frac{1}{g}\Phi\right) + \left(\partial_0 \overline{c}\right)\left(\partial_0 c\right)\Bigg]\end{aligned} \tag{31}$$

The last term in the above equation is the BRST-invariant gauge-fixing term. Proceeding classically, the Euler Lagrange equation for b reads:

$$-b = \left(g\partial_0 A_0 + g\partial_1 A_0 - g\partial_0 A_1\right) \tag{32}$$

which in turn (with the requirement $\hat{\delta} b = 0$) then implies:

$$\partial_0 \partial_0 c = 0 \tag{33}$$

The above equation is also an Euler-Lagrange equation (ELE) obtained by the variation of $\mathcal{L}_{BRST}$ with respect to $\overline{c}$. We define the bosonic momenta in the usual manner:

$$\Pi_0 := \frac{\partial}{\partial\left(\partial_0 A_0\right)}\mathcal{L}_{BRST} = gb \tag{34}$$

but for the fermionic momenta with directional derivatives we set

$$\Pi_c := \mathcal{L}_{BRST}\frac{\overleftarrow{\partial}}{\partial\left(\partial_0 c\right)} = \partial_0 \overline{c}; \quad \Pi_{\overline{c}} := \frac{\overrightarrow{\partial}}{\partial\left(\partial_0 \overline{c}\right)}\mathcal{L}_{BRST} = \partial_0 c \tag{35}$$

implying that the variable canonically conjugate to c is ($\partial_0 \overline{c}$) and the variable conjugate to $\overline{c}$ is ($\partial_0 c$). For writing the quantum Hamiltonian density from the Lagrangian density in the usual manner we remember that the former has to be Hermitian so that:

$$\begin{aligned}\mathcal{H}_{BRST} = \Bigg[& \Pi_u \partial_0 u + \Pi_v \partial_0 v + \hat{n}\cdot\hat{\Pi}_{\hat{n}} w + \frac{1}{2}E^2 + E\partial_1 A_0 - \lambda\left(\hat{n}^2 - 1\right) \\ & + \frac{1}{2}g^2 b^2 + \Pi_c \Pi_{\overline{c}} - gb\left(\partial_0 A_0 - \partial_0 A_1 + \partial_1 A_0\right)\Bigg]\end{aligned} \tag{36}$$

We can check the consistency of our definitions of the the fermionic momenta by looking at the Hamilton's equations for the fermionic variables:

$$\partial_0 c = \frac{\overrightarrow{\partial}}{\partial \Pi_c}\mathcal{H}_{BRST}; \quad \partial_0 \overline{c} = \mathcal{H}_{BRST}\frac{\overleftarrow{\partial}}{\partial \Pi_{\overline{c}}} \tag{37}$$

We thus see that

$$\partial_0 c = \frac{\overrightarrow{\partial}}{\partial \Pi_c}\mathcal{H}_{BRST} = \Pi_{\overline{c}}; \quad \partial_0 \overline{c} = \mathcal{H}_{BRST}\frac{\overleftarrow{\partial}}{\partial \Pi_{\overline{c}}} = \Pi_c \tag{38}$$

is in agreement with our definitions of the Fermionic momenta. Also, for the operators $c, \overline{c}, \partial_0 c$ and $\partial_0 \overline{c}$, one needs to satisfy the anti-commutation relations of $\partial_0 c$ with $\overline{c}$ or of $\partial_0 \overline{c}$ with c, but not of c, with $\overline{c}$. In general, c and $\overline{c}$ are independent canonical variables and one assumes that [25] [26] [27] [28]:

$$\left\{\Pi_c, \Pi_{\overline{c}}\right\} = \left\{\overline{c}, c\right\} = 0; \quad \partial_0\left\{\overline{c}, c\right\} = 0; \quad \left\{\partial_0 \overline{c}, c\right\} = (-1)\left\{\partial_0 c, \overline{c}\right\} \tag{39}$$

where $\{,\}$ means an anti-commutator. We thus see that the anti-commulators in the above equation are non-trivial and need to be fixed. In order to fix these, we demand that c satisfy the Heisenberg equation:

$$\left[c,\mathcal{H}_{BRST}\right]=i\partial_0 c \tag{40}$$

and using the property $c^2=\bar{c}^2=0$ one obtains

$$\left[c,\mathcal{H}_{BRST}\right]=\left\{\partial_0\bar{c},c\right\}\partial_0 c \tag{41}$$

The last three equations then imply :

$$\left\{\partial_0\bar{c},c\right\}=(-1)\left\{\partial_0 c,\bar{c}\right\}=i \tag{42}$$

Here the minus sign in the above equation is nontrivial and implies the existence of states with negative norm in the space of state vectors of the theory.

The BRST charge operator Q is the generator of the BRST transformations. It is nilpotent and satisfies $Q^2=0$. It mixes operators which satisfy Bose and Fermi statistics. According to its conventional definition, its commutators with Bose operators and its anti-commutators with Fermi operators for the present theory satisfy:

$$[\lambda,Q]=(-c),\ [A_0,Q]=\partial_0 c,\ [A_1,Q]=\partial_1 c \tag{43a}$$

$$[\hat{n},Q]=\hat{n}\partial_0 c,\ \left[\hat{\Pi}_{\hat{n}},Q\right]=\left(2\hat{n}c-\hat{\Pi}_{\hat{n}}\partial_0 c\right), \tag{43b}$$

$$\{\bar{c},Q\}=\left(\Pi_0+\Pi_\lambda+\hat{n}\cdot\hat{\Pi}_{\hat{n}}\right) \tag{43c}$$

$$\{\partial_0\bar{c},Q\}=(-1)\left(\partial_1 E+\hat{n}^2-1\right) \tag{43d}$$

All other commutators and anti-commutators of the theory involving Q and the other phase space variables of the theory are seen to vanish. In view of this, the BRST charge operator of the present theory could be written as:

$$Q=\int \mathrm{d}x^1\left[ic\left(\partial_1 E+\hat{n}^2-1\right)-i\partial_0 c\left(\Pi^0+\Pi_\lambda+\hat{n}\cdot\hat{\Pi}_{\hat{n}}\right)\right] \tag{44}$$

This equation implies that the set of states satisfying the conditions:

$$\Pi^0|\psi\rangle=0,\partial_1 E|\psi\rangle=0,\Pi_\lambda|\psi\rangle=0,\left(\hat{n}^2-1\right)|\psi\rangle=0,\left(\hat{n}\cdot\hat{\Pi}_{\hat{n}}\right)|\psi\rangle=0 \tag{45}$$

belong to the dynamically stable subspace of states $|\psi\rangle$ satisfying $Q|\psi\rangle=0$, *i.e.*, it belongs to the set of BRST-invariant states.

In order to understand the condition needed for recovering the physical states of the theory we rewrite the operators c and $\bar{c}$ in terms of fermionic annihilation and creation operators. For this purpose we consider Euler-Lagrange equation for the variable c derived earlier. The solution of this equation gives (for the instant-form time $x^0\equiv t$) the Heisenberg operators $c(t)$ and correspondingly $\bar{c}(t)$ in terms of the fermionic annihilation and creation operators as:

$$c(t)=G(t)+F(t),\quad \bar{c}(t)=G^\dagger(t)+F^\dagger(t) \tag{46}$$

which at the instant-form time $t=0$ imply

$$c\equiv c(0)=F,\quad \bar{c}(t)\equiv\bar{c}(0)=F^\dagger \tag{47a}$$

$$\partial_0 c(t) \equiv \partial_0 c(0) = G, \quad \partial_0 \bar{c}(t) \equiv \partial_0 \bar{c}(0) = G^\dagger \tag{47b}$$

By imposing the conditions (obtained earlier):

$$c^2 = \bar{c}^2 = \{\bar{c}, c\} = \{\partial_0 \bar{c}, \partial_0 c\} = 0 \tag{48a}$$

$$\{\partial_0 \bar{c}, c\} = (-1)\{\partial_0 c, \bar{c}\} = i \tag{48b}$$

we then obtain

$$F^2 = \left(F^\dagger\right)^2 = \left\{F^\dagger, F\right\} = \left\{G^\dagger, G\right\} = 0, \left\{G^\dagger, F\right\} = (-1)\left\{G, F^\dagger\right\} = i \tag{49}$$

Now let $|0\rangle$ denote the fermionic vacuum for which

$$G|0\rangle = F|0\rangle = 0 \tag{50}$$

Defining $|0\rangle$ to have norm one, the last three equations imply

$$\langle 0|FG^\dagger|0\rangle = i, \quad \langle 0|GF^\dagger|0\rangle = -i \tag{51}$$

so that

$$G^\dagger|0\rangle \neq 0, \quad F^\dagger|0\rangle \neq 0 \tag{52}$$

The theory is thus seen to possess negative norm states in the fermionic sector. The existence of these negative norm states as free states of the fermionic part of $\mathcal{H}_{BRST}$ is however, irrelevant to the existence of physical states in the orthogonal subspace of the Hilbert space. In terms of annihilation and creation operators $\mathcal{H}_{BRST}$ is:

$$\begin{aligned} \mathcal{H}_{BRST} = \Big[& \Pi_u \partial_0 u + \Pi_v \partial_0 v + \hat{n} \cdot \hat{\Pi}_{\hat{n}} w + \frac{1}{2} E^2 + E \partial_1 A_0 - \lambda\left(\hat{n}^2 - 1\right) \\ & + \frac{1}{2}\left(\Pi^0\right)^2 - \Pi^0\left(\partial_0 A_0 - \partial_0 A_1 + \partial_1 A_0\right) + G^\dagger G \Big] \end{aligned} \tag{53}$$

and the BRST charge operator of the present theory could be written as:

$$Q = \int \mathrm{d}x^1 \left[iF\left(\partial_1 E + \hat{n}^2 - 1\right) - iG\left(\Pi^0 + \Pi_\lambda + \hat{n} \cdot \hat{\Pi}_{\hat{n}}\right) \right] \tag{54}$$

Now because $Q|\psi\rangle = 0$, the set of states annihiliated by Q contains not only the set for which the constraints of the theory hold but also additional states for which

$$F|\psi\rangle = G|\psi\rangle = 0$$

$$\Pi^0|\psi\rangle \neq 0, \partial_1 E|\psi\rangle \neq 0, \Pi_\lambda|\psi\rangle \neq 0, \left(\hat{n}^2 - 1\right)|\psi\rangle \neq 0, \left(\hat{n} \cdot \hat{\Pi}_{\hat{n}}\right)|\psi\rangle \neq 0 \tag{55}$$

Now because $Q|\psi\rangle = 0$, the set of states annihilated by Q contains not only the set for which the constraints of the theory hold but also additional states for which the constraints of the theory do not hold. However in our considerations, the Hamiltonian is also invariant under the anti-BRST transformations given by:

$$\bar{\hat{\delta}}\lambda = -\frac{1}{g}\partial_0 \bar{c}, \bar{\hat{\delta}} A_0 = -\frac{1}{g}\partial_0 \bar{c}, \bar{\hat{\delta}} A_1 = -\frac{1}{g}\partial_1 \bar{c}, \bar{\hat{\delta}} u = -\frac{1}{g}\partial_0 \partial_0 \bar{c} \tag{56a}$$

$$\bar{\hat{\delta}}\hat{\Pi}_{\hat{n}} = \bar{\hat{\delta}}\Pi^0 = \bar{\hat{\delta}} E = \bar{\hat{\delta}}\Pi_\lambda = \bar{\hat{\delta}}\Pi_u = \bar{\hat{\delta}}\Pi_v = \bar{\hat{\delta}}\Pi_w = 0 \tag{56b}$$

$$\bar{\hat{\delta}}\hat{n} = \bar{\hat{\delta}}v = \bar{\hat{\delta}}w = 0, \bar{\hat{\delta}}\bar{c} = 0, \bar{\hat{\delta}}c = -b, \bar{\hat{\delta}}b = 0 \tag{56c}$$

with the generator or anti-BRST charge:

$$\bar{Q} = \int \mathrm{d}x^1 \left[-i\bar{c}\left(\partial_1 E + \hat{n}^2 - 1\right) + i\partial_0\bar{c}\left(\Pi^0 + \Pi_\lambda + \hat{n}\cdot\hat{\Pi}_{\hat{n}}\right)\right] \tag{57}$$

or

$$\bar{Q} = \int \mathrm{d}x^1 \left[-iF^\dagger\left(\partial_1 E + \hat{n}^2 - 1\right) + iG^\dagger\left(\Pi^0 + \Pi_\lambda + \hat{n}\cdot\hat{\Pi}_{\hat{n}}\right)\right] \tag{58}$$

We also have

$$\partial_0 Q = \left[Q, \mathcal{H}_{BRST}\right] = 0, \quad \partial_0\bar{Q} = \left[\bar{Q}, \mathcal{H}_{BRST}\right] = 0 \tag{59}$$

with

$$\mathcal{H}_{BRST} = \int \mathrm{d}x^1 \mathcal{H}_{BRST} \tag{60}$$

and we further impose the dual condition that both Q and $\bar{Q}$ annihilate physical states, implying that:

$$Q\left|\psi\right\rangle = 0 \quad \text{and} \quad \bar{Q}\left|\psi\right\rangle = 0 \tag{61}$$

The states for which the constraints of the theory hold, satisfy both of these conditions and are in fact, the only states satisfying both of these conditions, since with

$$G^\dagger G = (-1)GG^\dagger \tag{62}$$

there are no states of this operator with $G^\dagger\left|\psi\right\rangle = 0$ and $F^\dagger\left|\psi\right\rangle = 0$, and hence no free eigenstates of the fermionic part of $\mathcal{H}_{BRST}$ that are annihilated by each of G, $G^\dagger$, F, and $F^\dagger$. Thus the only states satisfying $Q\left|\psi\right\rangle = 0$ and $\bar{Q}\left|\psi\right\rangle = 0$ are those that satisfy the constraints of the theory.

Now because $Q\left|\psi\right\rangle = 0$, the set of states annihilated by Q contains not only the set of states for which the constraints of the theory hold but also additional states for which the constraints of the theory do not hold. This situation is, however, easily avoided by additionally imposing on the theory, the dual condition: $Q\left|\psi\right\rangle = 0$ and $\bar{Q}\left|\psi\right\rangle = 0$. By imposing both of these conditions on the theory simultaneously, one finds that the states for which the constraints of the theory hold are the only states satisfying both of these conditions. This is traced to the conditions on the fermionic variables c and $\bar{c}$ which constrain the solutions such that one cannot have simultaneously c, $\partial_0 c$ and $\bar{c}$, $\partial_0\bar{c}$, applied to $\left|\psi\right\rangle$ giving zero. Thus the only states satisfying $Q\left|\psi\right\rangle = 0$ and $\bar{Q}\left|\psi\right\rangle = 0$ are those that satisfy the constraints of the theory and they belong to the set of BRST-invariant as well as to the set of anti-BRST-invariant states.

Alternatively, one can understand the above point in terms of fermionic annihiliation and creation operators as follows. The condition $Q\left|\psi\right\rangle = 0$ implies the that the set of states annihilated by Q contains not only the states for which the constraints of the theory hold but also additional states for which the constraints do not hold. However, $\bar{Q}\left|\psi\right\rangle = 0$ guarantees that the set of states annihilated by $\bar{Q}$ contains only the states for which the constraints hold, simply because $G^\dagger\left|\psi\right\rangle \neq 0$ and $F^\dagger\left|\psi\right\rangle \neq 0$. Thus in this alternative way also, we see that the states satisfying $Q|\psi\rangle = \bar{Q}|\psi\rangle = 0$ are only those states which satisfy the constraints of the theory and we also see that these states belong to the set of BRST-invariant states as well as to the set of anti-BRST invariant states. This completes the BRST formulation of the theory.

5. Summary and Discussion

In the present work, we have considered the restricted gauge theory of quantum chromodynamics (QCD) in one-space one-time dimension (QCD_2) à la Cho *et al.* [1]-[14]. We have summarized the basics of the theory in Section 2 where the motivations of our present studies have also been discussed and are being omitted here the sake of brevity. The theory under our present investigation is seen to be GI and we have studied its quantization using Hamiltonian [24], path integral [25] [26] [27] [28] and Becchi-Rouet-Stora and Tyutin (BRST) [29] [30] [31], formulations [24]-[31], in the usual instant-form (IF) of dynamics (on the hyperplanes: $x^0 = t = \text{constant}$) [32] [33], under appropriate gauge-fixing conditions.

The restricted gauge theory of QCD_2 à la Cho *et al.* [1]-[14] and others [1]-[23] makes use of the so-called "Cho-decomposition", which is, in fact, the gauge independent decomposition of the non-Abelian potential into the restricted potential and the valence potential and it helps in the clarification of the topological structure of the non-Abelian gauge theory. This decomposition allows one to view QCD as the restricted gauge theory (made of the restricted potential) which is coupled to a gauge-covariant colored vector field (the valence potential). The restricted potential is defined in such a way that it allows a covariantly constant unit isovector $\hat{n}$ everywhere in space-time, which enables one to define the gauge-independent color direction everywhere in space-time and at the same time allows one to define the magnetic potential of the non-Abelian monopoles. It even has full SU(2) gauge degrees of freedom, in spite of the fact that it is restricted. Consequently, the restricted QCD made of the restricted potential describes a very interesting dual dynamics of its own, and plays a crucial role in the understanding of QCD. This restricted gauge theory of QCD is therefore very important and it is important to study its quantization using the standard constraint quantization methods, including the Hamiltonian, path integral and BRST quantizations, as we have done in the present work.

Acknowledgements

The authors UK and DSK thank James Vary for his warm hospitality at the Department of Physics and Astronomy, Iowa State University, Ames, Iowa, where a part of this work was carried out. This work was supported in part by the US Department of Energy under Grant No. DE-FG02-87ER40371.

Conflicts of Interest

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

References

[1] 't Hooft, G. (1981) *Nuclear Physics B*, **B190**, 455-478. https://doi.org/10.1016/0550-3213(81)90442-9

[2] Cho, Y.M. (2007) *Modern Physics Letters A*, **22**, 2799-2813. https://doi.org/10.1142/S0217732307025881

[3] Cho, Y.M. (2007) *Physics Letters B*, **644**, 208-211. https://doi.org/10.1016/j.physletb.2006.11.040

[4] Cho, Y.M. (2005) *Physics Letters B*, **616**, 101-107. https://doi.org/10.1016/j.physletb.2005.01.103

[5] Cho, Y.M. (2005) *Physical Review B*, **72**, 212516. https://doi.org/10.1103/PhysRevB.72.212516

[6] Cho, Y.M., Lee, H.W. and Park, D.G. (2002) *Physics Letters B*, **525**, 347-354. https://doi.org/10.1016/S0370-2693(01)01450-2

[7] Cho, Y.M. (2001) *Physical Review Letters*, **87**, 252001. https://doi.org/10.1103/PhysRevLett.87.252001

[8] Cho, Y.M. (2004) *Physics Letters B*, **603**, 88-93. https://doi.org/10.1016/j.physletb.2004.09.077

[9] Cho, Y.M., Kim, J.H. and Park, D.G. (2006) *Modern Physics Letters A*, **21**, 2789-2797. https://doi.org/10.1142/S0217732306021918

[10] Cho, Y.M., Walker, M.L. and Park, D.G. (2004) Monopole Condensation and Confinement of Color in SU(2) QCD. JHEP 0405, 073.

[11] Cho, Y.M. and Walker, M.L. (2004) *Modern Physics Letters A*, **19**, 2707-2716. https://doi.org/10.1142/S0217732304015750

[12] Cho, Y.M. and Pak, D.G. (2002) *Physical Review D*, **65**, 074027. https://doi.org/10.1103/PhysRevD.65.074027

[13] Bae, W.S., Cho, Y.M. and Zhang, P.M. (2008) *International Journal of Modern Physics A*, **23**, 267-298. https://doi.org/10.1142/S0217751X0803886X

[14] Bae, W.S., Cho, Y.M. and Kimm, S.-W. (2002) *Physical Review D*, **65**, 025005. https://doi.org/10.1103/PhysRevD.65.025005

[15] Faddeev, L. and Niemi, A. (1997) *Nature*, **387**, 58-61. https://doi.org/10.1038/387058a0

[16] Faddeev, L. and Niemi, A. (1999) *Physical Review Letters,* **82**, 1624-1627. https://doi.org/10.1103/PhysRevLett.82.1624

[17] Faddeev, L. and Niemi, A. (1999) *Physics Letters B,* **449**, 214-218. https://doi.org/10.1016/S0370-2693(99)00100-8

[18] Langman, E. and Niemi, A. (1999) *Physics Letters B,* **463**, 252-256. https://doi.org/10.1016/S0370-2693(99)01000-X

[19] Battye, R. and Sutcliffe, P.M. (1998) *Physical Review Letters,* **81**, 4798-4801. https://doi.org/10.1103/PhysRevLett.81.4798

[20] Shabanov, S. (1999) *Physics Letters B,* **458**, 322-330. https://doi.org/10.1016/S0370-2693(99)00612-7

[21] Shabanov, S. (1999) *Physics Letters B,* **463**, 263-272. https://doi.org/10.1016/S0370-2693(99)01024-2

[22] Gies, H. (2001) *Physical Review D,* **63**, Article ID: 025013.

[23] Kulshreshtha, D.S. (2015) *Few-Body Systems,* **56**, 565-569. https://doi.org/10.1007/s00601-015-1018-4

[24] Dirac, P.A.M. (1950) *Canadian Journal of Mathematics,* **2**, 129-148. https://doi.org/10.4153/CJM-1950-012-1

[25] Henneaux, M. and Teitelboim, C. (1976) *Annals Physics*, **100**, 227-261. https://doi.org/10.1016/0003-4916(76)90062-2

[26] Kulshreshtha, U. and Kulshreshtha, D.S. (2003) *Physics Letters B,* **555**, 255-263. https://doi.org/10.1016/S0370-2693(03)00056-X

[27] Kulshreshtha, U., Kulshreshtha, D.S. and Vary, J.P. (2015) *The European Physical Journal C*, **75**, 174. https://doi.org/10.1140/epjc/s10052-015-3377-x

[28] Kulshreshtha, U., Kulshreshtha, D.S. and Vary, J.P. (2016) *International Journal of Theoretical Physics*, **55**, 338-360. https://doi.org/10.1007/s10773-015-2665-4

[29] Becchi, C., Rouet, A. and Stora, A. (1974) *Physics Letters B,* **52**, 344-346. https://doi.org/10.1016/0370-2693(74)90058-6

[30] Tyutin, V. (1975) Lebedev Physics Institute Preprint-39.

[31] Nemeschansky, D., Preitschopf, C. and Weinstein, M. (1988) *Annals of Physics*, **183**, 226-268. https://doi.org/10.1016/0003-4916(88)90233-3

[32] Dirac, P.A.M. (1949) *Reviews of Modern Physics*, **21**, 392-399. https://doi.org/10.1103/RevModPhys.21.392

[33] Brodsky, S.J., Pauli, H.C. and Pinsky, S.S. (1998) *Physics Reports*, **301**, 299-486. https://doi.org/10.1016/S0370-1573(97)00089-6

Scientific Research Publishing

Journal of Modern Physics, 2018, 9, 2370-2377
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Vector Matrices Realization of Hurwitz Algebras

Daniel Sepunaru

RCQCE-Research Center for Quantum Communication, Holon Academic Institute of Technology, Holon, Israel
Email: danielsepunaru@walla.co.il

How to cite this paper: Sepunaru, D. (2018) A Vector Matrices Realization of Hurwitz Algebras. *Journal of Modern Physics*, **9**, 2370-2377.
https://doi.org/10.4236/jmp.2018.914151

Received: October 3, 2018
Accepted: December 2, 2018
Published: December 5, 2018

Abstract

We present here a realization of Hurwitz algebra in terms of 2 × 2 vector matrices which maintain the correspondence between the geometry of vector spaces that is used in the classical physics and the algebraic foundation underlying quantum theory. The multiplication rule we use is a modification of the one originally introduced by M. Zorn. We demonstrate that our multiplication is not intrinsically non-associative; the realization of the real and complex numbers is commutative and associative, the real quaternions maintain associativity and the real octonion matrices form an alternative algebra. Extension to the calculus of the matrices (with Hurwitz algebra valued matrix elements) of the arbitrary dimensions is straightforward. We briefly discuss applications of the obtained results to extensions of standard Hilbert space formulation in quantum physics and to alternative wave mechanical formulation of the classical field theory.

Keywords

Hilbert Spaces, Hurwitz Algebras, Zorn Multiplication

1. Introduction

While the mathematical formalism of classical physics is based on use of real vector spaces, quantum physics is typically formulated algebraically. Hence, a structure that allows for a connection between both these descriptions is necessary. Among the possible algebras relevant to this task, Hurwitz algebra plays a special role. It contains one-, two-, four- and eight-dimensional quadratic normal division algebras that form the only possible numerical systems. With Hurwitz algebra, we can generate the sequence of mathematical frameworks suitable for the description of dispersion-free [1] classical field theories as well as quantum field theories that obey Heisenberg dispersion relations that use Hilbert modules, which are functional analytical structures similar to the usual Hilbert spaces.

2. Matrix Treatment for Hurwitz Algebras

Consider the subsequence of those structures with real scalar products, all of whose dynamic variables are mutually commuting and whose states are real-, complex-, quaternion- and octonion-valued [2]:

$$(f,g)_R \equiv Tr(f,g) \tag{1}$$

where, for example, for a quaternion-valued f and g, (f,g) is quaternion-valued as well. The same structure may be alternatively generated by the four-dimensional vectors:

$$(f,g) = Tr(f,g) - e_1 Tr\{(f,g)e_1\} - e_2 Tr\{(f,g)e_2\} - e_3 Tr\{(f,g)e_3\} \tag{2}$$

$$-e(f,g)e_1 = Tr(f,g) - e_1 Tr\{(f,g)e_1\} + e_2 Tr\{(f,g)e_2\} + e_3 Tr\{(f,g)e_3\} \tag{3}$$

$$-e_2(f,g)e_2 = Tr(f,g) + e_1 Tr\{(f,g)e_1\} - e_2 Tr\{(f,g)e_2\} + e_3 Tr\{(f,g)e_3\} \tag{4}$$

$$-e_3(f,g)e_3 = Tr(f,g) + e_1 Tr\{(f,g)e_1\} + e_2 Tr\{(f,g)e_2\} - e_3 Tr\{(f,g)e_3\} \tag{5}$$

The sum of Equations (2), (3), (4) and (5) gives us:

$$(f,g)_R \equiv Tr(f,g) = \frac{1}{4}\left[(f,g) - e_1(f,g)e_1 - e_2(f,g)e_2 - e_3(f,g)e_3\right] \tag{6}$$

or in matrix notation

$$(\Psi,\Phi)_R = \frac{1}{4}\left[\bar{f}, -e_1\bar{f}, -e_2\bar{f}, -e_3\bar{f}\right] * \begin{bmatrix} g \\ ge_1 \\ ge_2 \\ ge_3 \end{bmatrix} \tag{7}$$

Similarly, the Hilbert module with a complex scalar product is generated by the sum of Equation (2) and Equation (3):

$$(f,g)_C \equiv Tr(f,g) - e_1 Tr\{(f,g)e_1\} = (f,g) - e_1(f,g)e_1 \tag{8}$$

In matrix notation

$$(\Psi,\Phi)_C = \frac{1}{2}\left[\bar{f}, -e_1\bar{f}\right] * \begin{bmatrix} g \\ ge_1 \end{bmatrix} \tag{9}$$

The Hilbert module with complex scalar products and octonion-valued states is generated in exactly the same manner. The usual Hilbert space obviously fits that procedure. This provides evidence of the existence of a uniform matrix treatment for all Hurwitz algebras.

First of all, let us consider 2 × 2 matrices. We have no difficulty in representing reals, complex and real quaternions, but the underlying Cayley-Dickson procedure prevents extending the 2 × 2 matrix to the 8-dimensional algebra of real octonions. In addition, the matrix obtained via Cayley-Dickson realization of real quaternions

$$q \Rightarrow \begin{pmatrix} q_0 - iq_3 & -iq_1 - q_2 \\ -iq_1 + q_2 & q_0 + iq_3 \end{pmatrix} \tag{10}$$

yields a physically erroneous mapping of space-time geometry

$$(x, y, z, t) \Rightarrow \begin{pmatrix} ct + z & x - iy \\ x + iy & ct - z \end{pmatrix} \tag{11}$$

since it violates the assumed isotropy of the space continuum. We, therefore modify the geometric vector matrix approach originally introduced by M. Zorn [3] [4] as follows:

1) For real numbers

$$X = x_0 \Rightarrow \begin{pmatrix} x_0 & 0 \\ 0 & x_0 \end{pmatrix} \tag{12}$$

2) For complex numbers

$$X = x_0 + x_1 i \equiv x_0 + \vec{x} \Rightarrow \begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix} = \begin{pmatrix} x_0 & x_1 i \\ x_1 i & x_0 \end{pmatrix} \tag{13}$$

3) For quaternions

$$X = x_0 + \sum_{i=1}^{3} x_i e_i \equiv x_0 + \vec{x} \Rightarrow \begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix} = \begin{pmatrix} x_0 & \sum_{i=1}^{3} x_i e_i \\ \sum_{i=1}^{3} x_i e_i & x_0 \end{pmatrix} \tag{14}$$

4) For octonions

$$X = x_0 + \sum_{i=1}^{7} x_i e_i \equiv x_0 + \vec{x} \Rightarrow \begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix} = \begin{pmatrix} x_0 & \sum_{i=1}^{7} x_i e_i \\ \sum_{i=1}^{7} x_i e_i & x_0 \end{pmatrix} \tag{15}$$

and the multiplication rule is defined by

$$\begin{aligned} Z = X \lozenge Y &\equiv \begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix} \lozenge \begin{pmatrix} y_0 & \vec{y} \\ \vec{y} & y_0 \end{pmatrix} \\ &= \begin{pmatrix} x_0 y_0 + \vec{x} \cdot \vec{y} & x_0 \vec{y} + y_0 \vec{x} + \vec{x} \times \vec{y} \\ x_0 \vec{y} + y_0 \vec{x} + \vec{x} \times \vec{y} & x_0 y_0 + \vec{x} \cdot \vec{y} \end{pmatrix} \end{aligned} \tag{16}$$

where

$$\begin{aligned} e_i \cdot e_j &= -\delta_{ij} \\ \vec{x} \cdot \vec{y} &= -x_i y_i = \vec{y} \cdot \vec{x} \\ \vec{x} \times \vec{y} &= \varepsilon_{ijk} x_i y_j e_k = -\vec{y} \times \vec{x}\,; \end{aligned} \tag{17}$$

ε_{ijk} are structural constants in the corresponding multiplication table (see **Appendix**). For quaternions this is usual a totally antisymmetric three-dimensional tensor; in the case of octonions it may also be considered as a Levi-Civita tensor in seven-dimensional space.

Explicitly, for quaternions we have

$$\vec{x} \times \vec{y} = (x_2 y_3 - x_3 y_2) e_1 + (x_3 y_1 - x_1 y_3) e_2 + (x_1 y_2 - x_2 y_1) e_3 \tag{18}$$

and for octonions

$$\begin{aligned}\vec{x}\times\vec{y} = {} & \left(x_2y_3 - x_3y_2\right)e_1 + \left(x_3y_1 - x_1y_3\right)e_2 + \left(x_1y_2 - x_2y_1\right)e_3 \\ & + \left(x_6y_5 - x_5y_6\right)e_1 + \left(x_6y_2 - x_2y_6\right)e_4 + \left(x_2y_5 - x_5y_2\right)e_7 \\ & + \left(x_4y_7 - x_7y_4\right)e_1 + \left(x_7y_2 - x_2y_7\right)e_5 + \left(x_2y_4 - x_4y_2\right)e_6 \\ & + \left(x_4y_6 - x_6y_4\right)e_2 + \left(x_1y_4 - x_4y_1\right)e_7 + \left(x_1y_6 - x_6y_1\right)e_5 \\ & + \left(x_5y_7 - x_7y_5\right)e_2 + \left(x_5y_1 - x_1y_5\right)e_6 + \left(x_7y_1 - x_1y_7\right)e_4 \\ & + \left(x_6y_7 - x_7y_6\right)e_3 + \left(x_7y_3 - x_3y_7\right)e_6 + \left(x_3y_6 - x_6y_3\right)e_7 \\ & + \left(x_5y_4 - x_4y_5\right)e_3 + \left(x_3y_5 - x_5y_3\right)e_4 + \left(x_4y_3 - x_3y_4\right)e_5\end{aligned} \tag{19}$$

Obviously,

$$\vec{x}\times\vec{x} = 0 \quad \text{and} \quad \vec{x}\cdot\vec{x} = -\sum_i x_i^2 \tag{20}$$

An involution is defined by

$$\overline{X} \equiv x_0 - \vec{x} \Rightarrow \begin{pmatrix} x_0 & -\vec{x} \\ -\vec{x} & x_0 \end{pmatrix} \tag{21}$$

and this satisfies the standard requirement

$$\overline{\overline{X}} = X$$

(which follows immediately from (21)).

$$\overline{X\Diamond Y} = \overline{Y}\Diamond\overline{X} \tag{22}$$

Proof:

$$\overline{X\Diamond Y} = \begin{pmatrix} x_0y_0 + \vec{x}\cdot\vec{y} & -x_0\vec{y} - y_0\vec{x} - \vec{x}\times\vec{y} \\ -x_0\vec{y} - y_0\vec{x} - \vec{x}\times\vec{y} & x_0y_0 + \vec{x}\cdot\vec{y} \end{pmatrix} \tag{23}$$

$$\begin{aligned}\overline{Y}\Diamond\overline{X} &= \begin{pmatrix} y_0 & -\vec{y} \\ -\vec{y} & y_0 \end{pmatrix}\Diamond\begin{pmatrix} x_0 & -\vec{x} \\ -\vec{x} & x_0 \end{pmatrix} \\ &= \begin{pmatrix} x_0y_0 + \vec{x}\cdot\vec{y} & -x_0\vec{y} - y_0\vec{x} - \vec{x}\times\vec{y} \\ -x_0\vec{y} - y_0\vec{x} - \vec{x}\times\vec{y} & x_0y_0 + \vec{x}\cdot\vec{y} \end{pmatrix}\end{aligned} \tag{24}$$

Now we are in the position to prove the following statement: The algebras defined by Equations (12), (13), (14), (15), (16) and (17) are quadratic normal division algebras.

Proof:

1) $Tr\left(X\right) \equiv X + \overline{X} = 2x_0$ (25)

2) $Det\left(X\right) \equiv N\left(X\right) \equiv X\Diamond\overline{X} = \begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix}\Diamond\begin{pmatrix} x_0 & -\vec{x} \\ -\vec{x} & x_0 \end{pmatrix} = \begin{pmatrix} x_0^2 - \vec{x}\cdot\vec{x} & 0 \\ 0 & x_0^2 - \vec{x}\cdot\vec{x} \end{pmatrix}$

Then

$$\begin{aligned}& X^2 - Tr\left(X\right)X + N\left(X\right) \\ & = \begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix}\Diamond\begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix} - 2x_0\begin{pmatrix} x_0 & \vec{x} \\ \vec{x} & x_0 \end{pmatrix} + \left(x_0^2 - \vec{x}\cdot\vec{x}\right)I = 0\end{aligned} \tag{26}$$

From the uniqueness of the Hurwitz algebras it follows that the realization discussed above has the following properties:

1) In one-dimensional algebra of reals and in two-dimensional algebra of complex numbers

$$X\Diamond Y = Y\Diamond X \quad \text{(commutative)}$$
$$X\Diamond\left(Y\Diamond Z\right) = \left(X\Diamond Y\right)\Diamond Z \quad \text{(associative)} \tag{27}$$

2) In four-dimensional algebra of real quaternions

$$X\Diamond\left(Y\Diamond Z\right) = \left(X\Diamond Y\right)\Diamond Z \quad \text{(associative)} \tag{28}$$

3) In eight-dimensional algebra of real octonions

$$X^2\Diamond Y = X\Diamond\left(X\Diamond Y\right) \quad \text{(left alternative)} \tag{29}$$

$$X\Diamond Y^2 = \left(X\Diamond Y\right)\Diamond Y \quad \text{(right alternative)} \tag{30}$$

Indeed, the validity the above statements may be demonstrated through direct matrix calculations. However, as they are rather cumbersome, we will only provide the useful relations for it:

1) All Hurwitz algebras hold

$$\vec{x}\cdot\vec{y} = \vec{y}\cdot\vec{x} \tag{31}$$

$$\vec{x}\times\vec{y} = -\vec{y}\times\vec{x} \tag{32}$$

$$\vec{x}\cdot\left(\vec{y}\times\vec{z}\right) = \vec{z}\cdot\left(\vec{x}\times\vec{y}\right) = \vec{y}\cdot\left(\vec{z}\times\vec{x}\right) \tag{33}$$

2) For quaternions

$$\vec{x}\times\left(\vec{y}\times\vec{z}\right) = \left(\vec{x}\cdot\vec{y}\right)\vec{z} - \left(\vec{x}\cdot\vec{z}\right)\vec{y} \tag{34}$$

Using relations (33) and (34) we have

$$\begin{aligned}
&\left(x\Diamond y\right)\Diamond z - x\Diamond\left(y\Diamond z\right)\\
&= \left[\left(\vec{x}\times\vec{y}\right)\cdot\vec{z} - \left(\vec{y}\times\vec{z}\right)\cdot\vec{x}\right] + \left[\left(\vec{x}\cdot\vec{y}\right)\vec{z} - \left(\vec{y}\cdot\vec{z}\right)\vec{x} + \left(\vec{x}\times\vec{y}\right)\times\vec{z} - \vec{x}\times\left(\vec{y}\times\vec{z}\right)\right]\\
&= 0
\end{aligned} \tag{35}$$

3) For octonions

$$\vec{x}\times\left(\vec{x}\times\vec{y}\right) = -\left(\vec{x}\cdot\vec{y}\right)\vec{x} + \left(\vec{x}\cdot\vec{x}\right)\vec{y} \tag{36}$$

Using (33) for the scalar component of the alternator we have

$$\left(\vec{x}\times\vec{y}\right)\cdot\vec{z} - \left(\vec{y}\times\vec{z}\right)\cdot\vec{x} = 0 \tag{37}$$

Therefore,

$$Tr\left[\left(x\Diamond y\right)\Diamond z\right] = Tr\left[x\Diamond\left(y\Diamond z\right)\right] \tag{38}$$

Thus, calculation of scalar products in the real Hilbert module with octonion-valued states may be performed neglecting their non-associativity. Obviously, we also have

$$Tr\left[\left(x\Diamond y\right)\Diamond z\right] = Tr\left[z\Diamond\left(x\Diamond y\right)\right] \tag{39}$$

We have obtained the properties of associativity and commutativity which are both needed to formulate a dispersion-free field theory [2].

A detailed discussion of self-adjoint operators (dynamic variables) in those frameworks will be presented in a separate publication.

Using (36) for the vector component of the alternator we have

$$\begin{aligned}&\left(\vec{x}\cdot\vec{x}\right)\vec{y}-\left(\vec{x}\cdot\vec{y}\right)\vec{x}-\vec{x}\times\left(\vec{x}\times\vec{y}\right)\\&=\left(\vec{x}\cdot\vec{x}\right)\vec{y}-\left(\vec{x}\cdot\vec{y}\right)\vec{x}+\left(\vec{x}\cdot\vec{y}\right)\vec{x}-\left(\vec{x}\cdot\vec{x}\right)\vec{y}=0\end{aligned} \tag{40}$$

or

$$x^2\Diamond y=x\Diamond\left(x\Diamond y\right)\quad\text{(left alternative)} \tag{41}$$

Similarly,

$$\begin{aligned}&\left(\vec{x}\cdot\vec{y}\right)\vec{y}-\left(\vec{y}\cdot\vec{y}\right)\vec{x}+\left(\vec{x}\times\vec{y}\right)\times\vec{y}\\&=\left(\vec{x}\cdot\vec{y}\right)\vec{y}-\left(\vec{y}\cdot\vec{y}\right)\vec{x}-\left(\vec{x}\cdot\vec{y}\right)\vec{y}+\left(\vec{y}\cdot\vec{y}\right)\vec{x}=0\end{aligned} \tag{42}$$

or

$$y\Diamond x^2=\left(y\Diamond x\right)\Diamond x\quad\text{(right alternative)} \tag{43}$$

Then the flexibility and the Moufang identities follow

$$\left(x\Diamond y\right)\Diamond x=x\Diamond\left(y\Diamond x\right) \tag{44}$$

$$\left(x\Diamond a\Diamond x\right)\Diamond y=x\Diamond\left[a\Diamond\left(x\Diamond y\right)\right] \tag{45}$$

$$y\Diamond\left(x\Diamond a\Diamond x\right)=\left[\left(y\Diamond x\right)\Diamond a\right]\Diamond x \tag{46}$$

$$\left(x\Diamond y\right)\Diamond\left(a\Diamond x\right)=x\Diamond\left(y\Diamond a\right)\Diamond x \tag{47}$$

Consider now matrices of arbitrary dimension with matrix elements belonging to one of the Hurwitz algebras. Then the product matrix is defined by the usual multiplication rule:

$$\begin{pmatrix}Z_{11}&Z_{12}&\cdots&Z_{1n}\\Z_{21}&Z_{22}&\cdots&Z_{2n}\\\vdots&\vdots&\ddots&\vdots\\Z_{n1}&Z_{n2}&\cdots&Z_{nn}\end{pmatrix}\equiv\begin{pmatrix}X_{11}&X_{12}&\cdots&X_{1n}\\X_{21}&X_{22}&\cdots&X_{2n}\\\vdots&\vdots&\ddots&\vdots\\X_{n1}&X_{n2}&\cdots&X_{nn}\end{pmatrix}*\begin{pmatrix}Y_{11}&Y_{12}&\cdots&Y_{1n}\\Y_{21}&Y_{22}&\cdots&Y_{2n}\\\vdots&\vdots&\ddots&\vdots\\Y_{n1}&Y_{n2}&\cdots&Y_{nn}\end{pmatrix} \tag{48}$$

$$Z_{ij}\equiv\sum_{k=1}^{n}X_{ik}Y_{kj};\quad i,j=1,2,\cdots,n \tag{49}$$

where

$$\begin{aligned}Z_{ij}&\equiv\begin{pmatrix}Z_{ij}^0&\vec{Z}_{ij}\\\vec{Z}_{ij}&Z_{ij}^0\end{pmatrix}=\sum_{k=1}^{n}\begin{pmatrix}x_{ik}^0&\vec{x}_{ik}\\\vec{x}_{ik}&x_{ik}^0\end{pmatrix}\Diamond\begin{pmatrix}y_{kj}^0&\vec{y}_{kj}\\\vec{y}_{kj}&y_{kj}^0\end{pmatrix}\\&\equiv\sum_{k=1}^{n}\begin{pmatrix}x_{ik}^0y_{kj}^0+\vec{x}_{ik}\cdot\vec{y}_{kj}&x_{ik}^0\vec{y}_{kj}+y_{kj}^0\vec{x}_{ik}+\vec{x}_{ik}\times\vec{y}_{kj}\\x_{ik}^0\vec{y}_{kj}+y_{kj}^0\vec{x}_{ik}+\vec{x}_{ik}\times\vec{y}_{kj}&x_{ik}^0y_{kj}^0+\vec{x}_{ik}\cdot\vec{y}_{kj}\end{pmatrix}\end{aligned} \tag{50}$$

$\forall X_{ik},Y_{ik},Z_{ik}$ elements of R, C, H and O algebras.

Thus, the product matrix is defined as the usual sum of pairs of multipliers and the product of each pair is defined by the vector multiplication introduced above. The trace and determinant of the product matrix are always real and are defined according to the usual rules. For example,

$$Tr\left(Z\right)=\sum_{i=1}^{n}Z_{ii}^0 \tag{51}$$

Therefore, the result of the calculation is unambiguous.

3. Conclusions

Having discussed the geometric extension of conventional matrix multiplication which is uniformly valid for all quadratic normal division algebras, I would like, in conclusion, to emphasize that the suggested matrix realization is of crucial importance for quaternion and octonion extensions of standard functional analysis since the real as well as the complex Hilbert modules require the use of multicomponent states. The results obtained allow for the introduction and investigation of the operators necessary for the description of the system dynamics as well as for the observables (self-adjoint operators) [5] [6]. In addition, the transition from the vector matrix to the standard one may provide an alternative mechanism for spontaneous breakdown of internal symmetries as suggested by the comparison of Equations (10) and (14). Historically, the multiplication operation over real numbers was first extended to physically relevant three-dimensional space and only later to spaces of arbitrary dimensions and signatures [7]. The invention of scalar matrix multiplication was an alternative to this generalization. It seems reasonable to expect that the vector matrix multiplication suggested here may be extended to additional types of algebras (Clifford, Lie, Jordan, etc.), but that lies outside the scope of this investigation.

Conflicts of Interest

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

References

[1] von Neumann, J. (1931) Mathematische Grundlagen der Quantenmechanik. Springer, Berlin.

[2] Sepunaru, D. (2006) quant-ph/0606121.

[3] Zorn, M. (1930) *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, **8**, 123-147.

[4] Zorn, M. (1933) *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, **9**, 395-402.

[5] Dirac, P.A.M. (1958) The Principles of Quantum Mechanics. 4th Edition, Clarendon Press, Oxford.

[6] Doughty, N.A. (1990) Lagrangian Interaction. Addison-Wesley, Boston.

[7] Cartan, E. (1938) Lecons sur la theorie des spineuers. *Actualitesscientifiques et industrielles*, Hermann, Paris, 643, 701.

Appendix

For readers who would like to verify the statements in this paper by direct calculation, I reproduce here the multiplication tables of Hurwitz algebra.

Table A1. Complex numbers.

	e_0	e_1
e_0	e_0	e_1
e_1	e_1	$-e_0$

Table A2. Quaternions.

	e_0	e_1	e_2	e_3
e_0	e_0	e_1	e_2	e_3
e_1	e_1	$-e_0$	e_3	$-e_2$
e_2	e_2	$-e_3$	$-e_0$	e_1
e_3	e_3	e_2	$-e_1$	$-e_0$

Table A3. Octonions.

	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7
e_0	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7
e_1	e_1	$-e_0$	e_3	$-e_2$	e_7	$-e_6$	e_5	$-e_4$
e_2	e_2	$-e_3$	$-e_0$	e_1	e_6	e_7	$-e_4$	$-e_5$
e_3	e_3	e_2	$-e_1$	$-e_0$	$-e_5$	e_4	e_7	$-e_6$
e_4	e_4	$-e_7$	$-e_6$	e_5	$-e_0$	$-e_3$	e_2	e_1
e_5	e_5	e_6	$-e_7$	$-e_4$	e_3	$-e_0$	$-e_1$	e_2
e_6	e_6	$-e_5$	e_4	$-e_7$	$-e_2$	e_1	$-e_0$	e_3
e_7	e_7	e_4	e_5	e_6	$-e_1$	$-e_2$	$-e_3$	$-e_0$

Journal of Modern Physics, 2018, 9, 2378-2390
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

On the Problem of Eigenschaften in Quantum and Classical Mechanics

Daniel Sepunaru

RCQCE-Research Center for Quantum Communication, Holon Academic Institute of Technology, Holon, Israel
Email: danielsepunaru@walla.co.il

How to cite this paper: Sepunaru, D. (2018) On the Problem of Eigenschaften in Quantum and Classical Mechanics. *Journal of Modern Physics*, **9**, 2378-2390.
https://doi.org/10.4236/jmp.2018.914152

Received: October 3, 2018
Accepted: December 2, 2018
Published: December 5, 2018

Abstract

We argue that in contrast to the classical physics, measurements in quantum mechanics should provide simultaneous information about all relevant relative amplitudes (pure states and the transitions between them) and all relevant relative phases. Simultaneity is needed since measurement changes the state of the system (in both quantum and in classical physics). We call that measurement procedure holographic detection. Mathematically, it is described by a set of mutually commuting selfadjoint operators that are similar and closely related to projections. We present explicit examples and discuss general features of the corresponding experimental setup which we identify as the quantum reference frame.

Keywords

Hilbert Spaces, Holographic Detection, Quantum Reference Frames

1. Introduction

Debates about the connection between hidden laws of nature and our ability to extract the information necessary to formulate these laws have a long history, perhaps as long as study of physics itself. This paper, while not related to the philosophical or metaphysical aspects of those discussions, puts forth certain point of view without intention to defend it or to convince the reader that it is only possible approach. We simply present how the process of knowledge acquisition is realized within that approach. We explore the analogy to the structure of field theories (classical electrodynamics, general relativity and non-relativistic quantum mechanics) and make a distinction between unobservable kinematical quantities which characterize a physical system and the measurable variables which define its dynamics. Since the main distinction between classical and quantum physics is the presence of new kinematic quantities—phases—we need to know how to measure the corresponding phase differences. We demonstrate that this measurement may be obtained by using a special experimental arrangement that we call quantum reference frames. This allows for communicating the required hidden unobservable information to the instruments of the observer. This simultaneously explains why the elementary unit of communication is given in terms of an indivisible bit.

The notion of the eigenschaften operator was first introduced by J. von Neumann [1] as a necessary ingredient of his theory of measurements. He suggested assigning that role to projection operators which define not only the space of quantum mechanical states but also the structure of that space and its complete, orthonormal basis. In our model, it is logically consistent to use eigenschaften operators that closely relate to projection operators but act on the whole space without distortion; that is, eigenschaften operators that are unitary.

The main feature of the measurement process is that measurement devices are macroscopic, obeying the laws of classical physics, whereas the systems being tested belong to the microscopic world and behave quantum mechanically. Indeed, the measurement setup should assure that the results obtained represent objective properties of the physical system being investigated and not the subjective imagination of the observer. Using classical physics, we complete that task by introducing reference frames such that the location of the detector defines both the frame's origin and the set of auxiliary macroscopic devices. This allows for establishment of a connection between frames that are separated by a finite space-time interval (comparison of the empirical data obtained must always be performed by the same observer). Similarly, in order to measure the relevant quantum dynamical variable a set of auxiliary macroscopic devices should be included in the classical setup to produce the necessary beam-splitting. Then the required phase differences can be measured in the usual way. This setup and recording procedure may be viewed as general holographic detection.

The organization of this paper is as follows:

Section 2 presents a discussion of the relevant kinematics of the quantum theory. Section 3 introduces the unitary, self-adjoint operators which we identify as adequate eigenschaften operators.

Section 4 discusses the quantum frames of reference making a close analogy to the inertial frames of classical physics.

2. The Kinematics of Quantum Mechanical Theory

We restrict ourselves to discussion of single particle states, avoiding complications introduced by special relativity. We use an orthodox kinematic approach based on the mathematical framework of Hilbert metric spaces. That means that we assume that there exists at least one self-adjoin operator that generates this space. That operator is supposed to describe the dynamics of a single particle that is completely isolated from the external world. All measurable quantities are also described by self-adjoint operators. In particular, projection operators, density matrices, etc. are treated as special kinds of observables, whereas the fundamental quantity associated with the state of the physical system is treated as a wave function [2]. In contrast to operators that are geometric transformations of the given vector space, wave functions are vectors that form that space and are both unobservable and incapable of being measured directly, at least in principle.

The transition from the sterile situation of a single isolated particle to the real-life physical system is achieved through introduction of the local interactions of the test particle with the fields generated by the rest of the external world. These interactions are introduced using the principle of local gauge invariance. The required complexity emerges from the statistical nature of the environment. This approach is identical to the conventional one that has long been established in the development of classical physics over the centuries except that the definition of (fundamental) interactions is now connected to the new physics, since we are dealing with matter waves.

The fundamental property of the quantum mechanical states is as expressed in terms of the linear superposition principle is:

If $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are two different states of the system, then

$$|\Psi\rangle = a|\Psi_1\rangle + b|\Psi_2\rangle \tag{1}$$

is also a state of the system. Equivalently, we may write:

$$|\Psi\rangle = a\begin{pmatrix} |\Psi_1\rangle \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ |\Psi_2\rangle \end{pmatrix} \tag{2}$$

or

$$|\Psi\rangle = \begin{pmatrix} a|\Psi_1\rangle \\ b|\Psi_2\rangle \end{pmatrix}; \ \langle\Psi_1|\Psi_1\rangle = \langle\Psi_2|\Psi_2\rangle = 1; \ \langle\Psi_1|\Psi_2\rangle = 0 \tag{3}$$

However, that seemingly innocent-looking mathematical expressions leads to a dramatic change in the physics of the described system, since the presence of the second orthogonal component is the necessary and sufficient condition that now the above function describes the extended object:

Theorem [3]: if $\hat{A}^{+} = \hat{A}$ and $\langle\Psi_1|\Psi_2\rangle = 0$; $\langle\Psi_1|\Psi_1\rangle = \langle\Psi_2|\Psi_2\rangle = 1$;

Then

$$\hat{A}|\Psi_1\rangle = a|\Psi_1\rangle + b|\Psi_2\rangle \tag{4}$$

$$a = \langle\Psi_1|\hat{A}|\Psi_1\rangle = \langle\hat{A}\rangle \equiv \bar{A} \tag{5}$$

$$|b|^2 = bb^* = \langle\Psi_1|\left(\hat{A}^2 - a^2\right)|\Psi_1\rangle \equiv (\Delta A)^2 \tag{6}$$

can be decomposed.

Proof:

$$\langle\Psi_1|\left(\hat{A}^2 - a^2\right)|\Psi_1\rangle = \langle\Psi_1|\hat{A}^2|\Psi_1\rangle - a^2 \ =$$

$$\left(a\langle\Psi_1| + b^*\langle\Psi_2|\right)\left(a|\Psi_1\rangle + b|\Psi_2\rangle\right) - a^2 = bb^*$$

Therefore, what we need to reconstruct in the properly performed quantum mechanical measurement is a picture. Since equations of motion are intrinsically complex, the quantum mechanical system must be described by a two-component state function at least, due to the Euler relation:

$$\exp(i\varphi) = \cos\varphi + i\sin\varphi .$$

In contrast to classical physics, quantum mechanics is the physics of extended objects; it is the theory of matter fields. Now, due to D. Hilbert's spectral decomposition theorem [4], any $\hat{A}$, such that $\hat{A} = \hat{A}^{+}$ may be expressed in terms of one-dimensional projectors:

$$\hat{A} = \sum_{n} \lambda_n \hat{P}_n \tag{7}$$

where

$$\hat{P}_n^{+} = \hat{P}_n ; \hat{P}_n \hat{P}_m = \delta_{nm} \hat{P}_m ; \sum \hat{P}_n = \hat{I} ; \tag{8}$$

or, in Dirac notation:

$$\hat{P}_n = |\varphi_n\rangle\langle\varphi_n| ; \tag{9}$$

λ_n are eigenvalues of operator $\hat{A}$ and $|\varphi_n\rangle$ are its eigenfunctions. The set of eigenfunctions forms a complete orthonormal basis. Thus, the space obtained is the metric space suitable for physical applications, hence, Operator (9) defines a pure state. More generally, one introduces the density matrix

$$\begin{gathered} \hat{\rho} = |\varphi\rangle\langle\varphi| \\ \rho_{ij} = \langle\varphi_i|\hat{\rho}|\varphi_j\rangle = \langle\varphi_i|\varphi\rangle\langle\varphi|\varphi_j\rangle \\ \rho_{ij} = \langle\varphi_i|\varphi\rangle\langle\varphi_j|\varphi\rangle^{*} \end{gathered} \tag{10}$$

or

$$\hat{\rho} = \sum_{n} w_n |\varphi_n\rangle\langle\varphi_n| \tag{11}$$

We may try to use linear algebra in order to clarify the difference between uni- and multi-component states. Using Heisenberg-Schrödinger notation, we may write:

$$\begin{gathered} \hat{P}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} ; \ \hat{P}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ \hat{P}_1 + \hat{P}_2 = \hat{I} \end{gathered} \tag{12}$$

We consider the two-component case only for its simplicity, generalization to the non-generate finite dimension case is straightforward:

Now consider the two-component wave function. Then

$$\begin{gathered} |\Psi\rangle = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \\ aa^{*} + bb^{*} = 1 \end{gathered} \tag{13}$$

The corresponding density matrix

$$\hat{\rho} = \begin{pmatrix} aa^* & ab^* \\ ba^* & bb^* \end{pmatrix} \tag{14}$$

may be obtained using Kronecker product multiplication

$$\hat{\rho} = \begin{pmatrix} a \\ b \end{pmatrix} \otimes \left(a^*, b^* \right) \tag{15}$$

However, equation (14) still describes a pure state, since

$$\hat{\rho} = \hat{\rho}^+ ; \quad \hat{\rho}^2 = \hat{\rho} ; \quad Tr\hat{\rho} = 1 \tag{16}$$

Let us introduce the notation

$$\hat{\hat{\rho}} = \begin{pmatrix} aa^* & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & bb^* \end{pmatrix} \tag{17}$$

Then,

$$\hat{\rho} = \hat{\hat{\rho}} + \begin{pmatrix} 0 & ab^* \\ ba^* & 0 \end{pmatrix} \tag{18}$$

Obviously,

$$\hat{\hat{\rho}} = \hat{\hat{\rho}}^+ ; \quad Tr\hat{\hat{\rho}} = 1 \tag{19}$$

But

$$\hat{\hat{\rho}}^2 \neq \hat{\hat{\rho}} \quad \text{if} \quad ab \neq 0 \tag{20}$$

$\hat{\hat{\rho}}$ is a mixture of two pure one-particle states $\begin{pmatrix} a \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ b \end{pmatrix}$. Clearly, this cannot be treated as a single particle state. In order to demonstrate this let us calculate the dispersion of the projection operator $\hat{\rho}$:

$$\langle \hat{\rho} \rangle \equiv Tr\left(\hat{\rho}\hat{\hat{\rho}} \right) = \left(aa^* \right)^2 + \left(bb^* \right)^2 = 1 - 2\left(aa^* \right)\left(bb^* \right). \tag{21}$$

If $ab \neq 0$, then

$$1 > \langle \hat{\rho} \rangle > 0 \quad \text{and} \quad 1 > \langle \hat{\rho} \rangle^2 > 0$$

Therefore,

$$\left(\Delta\hat{\rho} \right)^2 \equiv Tr\left(\hat{\rho}^2 \hat{\hat{\rho}} \right) - \left(Tr\left(\hat{\rho}\hat{\hat{\rho}} \right) \right)^2 = Tr\left(\hat{\rho}\hat{\hat{\rho}} \right)\left[1 - Tr\left(\hat{\rho}\hat{\hat{\rho}} \right) \right] > 0 \tag{22}$$

which contradicts the spectral decomposition theorem. Hence, the system state in our example is a pure state.

The operator $\hat{\rho}$ (Equation (14)) preserves the clear geometrical meaning of a one-dimensional, dispersion-free projector. If one starts with a well-defined reference frame, the complete set of those projectors allows the rotation of that new reference frame to the new axes. However, that set does not allow the extraction of information about the dispersions contained in the measurements of the transition amplitudes. The next section discusses the self-adjoint operators that allow for doing just that.

3. Eigenschaften Operators

From the logical point of view, it is natural to expect that projection operators do not provide an adequate means for obtaining information about all possible alternatives, since they destroy the orthogonal subspace of the Hilbert space: A true eigenschaften operator must be unitary. Together with the requirement of being observable ($\hat{H}^{+}=\hat{H}$), that leads to the following theorem:

Theorem:

If $\hat{H}^{+}=\hat{H}^{-1}$ (unitary) and $\hat{H}^{+}=\hat{H}$ (self-adjoint),

Then

$$\hat{H}^{2}=\hat{I}\,.$$

Proof: (23)

1) Suppose

$$\hat{H}^{+}=\hat{H}=\hat{H}^{-1},$$

then

$$\hat{H}\cdot\hat{H}=\hat{H}\cdot\hat{H}^{-1}=\hat{I}\,.$$

2) Suppose

$$\hat{H}^{2}=\hat{I}\quad\text{and}\quad\hat{H}^{+}=\hat{H}^{-1},$$

then

$$\hat{H}^{+}=\hat{H}\,.$$

From $\hat{H}^{2}=\hat{I}$ we have

$$\left(\hat{H}-\hat{I}\right)\cdot\left(\hat{H}+\hat{I}\right)=0 \tag{24}$$

Let us first consider the two-dimensional case. From Equation (24)

$$\lambda_1=1;\ \ \lambda_2=-1 \tag{25}$$

and due to the spectral composition theorem, we have

$$\hat{H}_2=\hat{P}_1-\hat{P}_2 \tag{26}$$

Since

$$\hat{P}_1+\hat{P}_2=\hat{I} \tag{27}$$

we finally obtain

$$\begin{aligned}\hat{P}_1&=\frac{\hat{I}+\hat{H}_2}{2}\\ \hat{P}_2&=\frac{\hat{I}-\hat{H}_2}{2}\end{aligned} \tag{28}$$

Now in terms of matrix mechanics we have

$$\begin{aligned}\hat{H}_2\left|\Psi_1\right\rangle&=\alpha_1\left|\Psi_1\right\rangle+\beta\mathrm{e}^{i\Delta\varphi}\left|\Psi_2\right\rangle\equiv\left|\Psi_3\right\rangle\\ \hat{H}_2\left|\Psi_2\right\rangle&=\beta\mathrm{e}^{-i\Delta\varphi}\left|\Psi_1\right\rangle+\alpha_2\left|\Psi_2\right\rangle\equiv\left|\Psi_4\right\rangle\end{aligned} \tag{29}$$

with

$$\begin{gathered}\langle\Psi_1|\Psi_1\rangle=\langle\Psi_2|\Psi_2\rangle=\langle\Psi_3|\Psi_3\rangle=\langle\Psi_4|\Psi_4\rangle=1\\ \langle\Psi_1|\Psi_2\rangle=\langle\Psi_3|\Psi_4\rangle=0\end{gathered}\tag{30}$$

Then,

$$\begin{aligned}&\beta\cdot(\alpha_1+\alpha_2)=0\\&\alpha_1^2+\beta^2=1\\&\alpha_2^2+\beta^2=1\end{aligned}\tag{31}$$

Since we are discussing here the measurement of the relevant parameters of quantum mechanical systems with non-vanishing dispersion, we consider only the $\beta\neq 0$ case. Then,

$$\alpha_1=-\alpha_2\equiv\alpha\tag{32}$$

or

$$Tr\left(\hat{H}_2\right)=0\tag{33}$$

The matrix elements

$$\langle\Psi_1|\hat{H}_2|\Psi_1\rangle=-\langle\Psi_2|\hat{H}_2|\Psi_2\rangle=\alpha\tag{34}$$

and

$$\langle\Psi_2|\hat{H}_2|\Psi_1\rangle=\beta e^{i\cdot\Delta\varphi}\tag{35}$$

are all we need to know about the quantum state. Both are measurable, $\left(\hat{H}_2\right)_{11}$ defines the spectrum and $\left(\hat{H}_2\right)_{12}$ defines the dispersion. The basis introduced above $|\Psi_3\rangle$ and $|\Psi_4\rangle$ is distinguished by the fact that it allows for simultaneous measurement of both spectrum and dispersion. The example of a two-level system should make this even clearer:

$$\begin{aligned}\hat{H}_2\left(e^{-i\cdot\omega_1 t}|\Psi_1\rangle\right)&=\alpha_1 e^{-i\cdot\omega_1 t}|\Psi_1\rangle+\beta e^{-i\cdot\omega_2 t}|\Psi_2\rangle\\ \hat{H}_2\left(e^{-i\cdot\omega_2 t}|\Psi_2\rangle\right)&=\beta e^{-i\cdot\omega_1 t}|\Psi_1\rangle-\alpha e^{-i\cdot\omega_2 t}|\Psi_2\rangle\end{aligned}\tag{36}$$

Then dropping the overall phase factor, we obtain

$$\begin{aligned}\hat{H}_2|\Psi_1\rangle&=\alpha|\Psi_1\rangle+\beta e^{i\cdot(\omega_1-\omega_2)t}|\Psi_2\rangle\\ \hat{H}_2|\Psi_2\rangle&=\beta e^{-i\cdot(\omega_1-\omega_2)t}|\Psi_1\rangle-\alpha|\Psi_2\rangle\end{aligned}\tag{37}$$

Using the relations in (31) we obtain the most general solution:

$$\hat{H}_2=\begin{pmatrix}\cos\gamma & e^{-i\cdot\Delta\varphi}\cdot\sin\gamma\\ e^{i\cdot\Delta\varphi}\cdot\sin\gamma & -\cos\gamma\end{pmatrix}\tag{38}$$

In particular, for $\Delta\varphi=0$ and $\gamma=45^\circ$ we obtain the Hadamard matrix of lowest order ($N=2$)

$$\hat{H}_2=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\tag{39}$$

that is well-known in image processing applications.

Now we demonstrate that the solution obtained is intrinsically consistent with the general statement [3] referred to above. The density matrix in our example

(Equation (14)) is

$$\hat{\rho} = \begin{pmatrix} \cos^2\gamma & \sin\gamma\cdot\cos\gamma\cdot \mathrm{e}^{-i\Delta\varphi} \\ \sin\gamma\cdot\cos\gamma\cdot \mathrm{e}^{i\Delta\varphi} & \sin^2\gamma \end{pmatrix} \tag{40}$$

Then

$$\left\langle \hat{H}_2 \right\rangle \equiv Tr\left(\hat{H}_2\hat{\rho}\right) = Tr\begin{pmatrix} \cos\gamma & \sin\gamma\cdot \mathrm{e}^{-i\Delta\varphi} \\ 0 & 0 \end{pmatrix} = \cos\gamma = \alpha \tag{41}$$

$$\left\langle \hat{H}_2^2 \right\rangle \equiv Tr\left(\hat{H}_2^2\hat{\rho}\right) = Tr\,\hat{\rho} = 1 \tag{42}$$

and

$$\left(\Delta\hat{H}_2\right)^2 \equiv \left\langle \hat{H}_2^2 \right\rangle - \left(\left\langle \hat{H}_2 \right\rangle\right)^2 = 1-\cos^2\gamma = \sin^2\gamma = \beta\beta^* \tag{43}$$

Consider now the three-component case (an analog to three-level quantum mechanical systems).

We prefer to explicitly discuss the three-component and the four-component cases, rather than the general n-dimensional situation which follows directly from the results obtained.

We have

$$\begin{aligned} \hat{H}_3\left|\Psi_1\right\rangle &= \alpha_1\left|\Psi_1\right\rangle + \beta e^{i\Delta\phi_1}\left|\Psi_2\right\rangle + \gamma \mathrm{e}^{i\Delta\varphi_2}\left|\Psi_3\right\rangle \equiv \left|\Psi_4\right\rangle \\ \hat{H}_3\left|\Psi_2\right\rangle &= \beta \mathrm{e}^{-i\Delta\varphi_1}\left|\Psi_1\right\rangle + \alpha_2\left|\Psi_2\right\rangle + \mu \mathrm{e}^{i\Delta\varphi_3}\left|\Psi_3\right\rangle \equiv \left|\Psi_5\right\rangle \\ \hat{H}_3\left|\Psi_3\right\rangle &= \gamma \mathrm{e}^{-i\Delta\varphi_2}\left|\Psi_1\right\rangle + \mu \mathrm{e}^{-i\Delta\varphi_3}\left|\Psi_2\right\rangle + \alpha_3\left|\Psi_3\right\rangle \equiv \left|\Psi_6\right\rangle \end{aligned} \tag{44}$$

with

$$\begin{aligned} &\left\langle \Psi_1 \middle| \Psi_1 \right\rangle = \left\langle \Psi_2 \middle| \Psi_2 \right\rangle = \left\langle \Psi_3 \middle| \Psi_3 \right\rangle = 1 \\ &\left\langle \Psi_4 \middle| \Psi_4 \right\rangle = \left\langle \Psi_5 \middle| \Psi_5 \right\rangle = \left\langle \Psi_6 \middle| \Psi_6 \right\rangle = 1 \\ &\left\langle \Psi_1 \middle| \Psi_2 \right\rangle = \left\langle \Psi_1 \middle| \Psi_3 \right\rangle = \left\langle \Psi_2 \middle| \Psi_3 \right\rangle = 0 \\ &\left\langle \Psi_4 \middle| \Psi_5 \right\rangle = \left\langle \Psi_4 \middle| \Psi_6 \right\rangle = \left\langle \Psi_5 \middle| \Psi_6 \right\rangle = 0 \end{aligned} \tag{45}$$

Then, the matrix elements of $\hat{H}_3$ are connected by the following relations:

$$\begin{aligned} &\left(Tr\left(\hat{H}_3\right)\right)^2 = \left(\alpha_1+\alpha_2+\alpha_3\right)^2 = 1 \\ &Tr\left(\hat{H}_3\right) = \pm 1 \\ &\beta^2 = \left(1\mp\alpha_1\right)\left(1\mp\alpha_2\right) \\ &\gamma^2 = \left(1\mp\alpha_1\right)\left(1\mp\alpha_3\right) \\ &\mu^2 = \left(1\mp\alpha_2\right)\left(1\mp\alpha_3\right) \\ &\Delta\varphi_3 = \Delta\varphi_2 - \Delta\varphi_1 \end{aligned} \tag{46}$$

Let us establish the connection between the eigenschaften and the projection operators here. Consider the uni-dimensional projection operators

$$\begin{aligned} &\hat{P}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \hat{P}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \hat{P}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \\ &\hat{P}_1 + \hat{P}_2 + \hat{P}_3 = \hat{I} \end{aligned} \tag{47}$$

Again, the most general one-dimensional projector may be written in the form

$$\hat{\rho} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \otimes \left(a^*, b^*, c^* \right) \qquad (48)$$
$$\hat{\rho}^+ = \hat{\rho}; \hat{\rho}^2 = \hat{\rho}; Tr\,\hat{\rho} = aa^* + bb^* + cc^* = 1$$

Then using spectral decomposition

$$\begin{aligned} &\hat{H}_3 = \lambda_1 \hat{P}_1 + \lambda_2 \hat{P}_2 + \lambda_3 \hat{P}_3 \\ &\hat{H}_3^2 = \hat{I} \end{aligned} \qquad (49)$$

we have

$$\begin{aligned} &\hat{H}_3^{(1)} = -\hat{P}_1 + \hat{P}_2 + \hat{P}_3 \\ &\hat{H}_3^{(2)} = \hat{P}_1 - \hat{P}_2 + \hat{P}_3 \\ &\hat{H}_3^{(3)} = \hat{P}_1 + \hat{P}_2 - \hat{P}_3 \end{aligned} \qquad (50)$$

Thus, we obtain

$$\begin{aligned} &\hat{P}_1 = \frac{\hat{I} - \hat{H}_3^{(1)}}{2} \\ &\hat{P}_2 = \frac{\hat{I} - \hat{H}_3^{(2)}}{2} \\ &\hat{P}_3 = \frac{\hat{I} - \hat{H}_3^{(3)}}{2} \end{aligned} \qquad (51)$$

However, only two equations are linearly independent

$$\hat{H}_3^{(1)} + \hat{H}_3^{(2)} + \hat{H}_3^{(3)} = \hat{I} \qquad (52)$$

and form the following commutative algebra

$$\begin{aligned} &\hat{H}_3^{(1)} \cdot \hat{H}_3^{(2)} = -\hat{H}_3^{(3)} \\ &\left[\hat{H}_3^{(i)}, \hat{H}_3^{(j)} \right] = 0; i, j = 1, 2, 3 \end{aligned} \qquad (53)$$

We conclude with a demonstration of the four-component case. The $\hat{H}_4$ operators ($\hat{H}_4^+ = \hat{H}_4$ and $\hat{H}_4^2 = \hat{I}$) have the form:

$$\hat{H}_4 = \begin{pmatrix} \alpha_1 & \beta e^{-i\Delta\varphi_1} & \gamma e^{-i\Delta\varphi_2} & \delta e^{-i\Delta\varphi_4} \\ \beta e^{i\Delta\varphi_1} & \alpha_2 & \mu e^{-i\Delta\varphi_3} & \upsilon e^{-i\Delta\varphi_5} \\ \gamma e^{i\Delta\varphi_2} & \mu e^{i\Delta\varphi_3} & \alpha_3 & \zeta e^{-i\Delta\varphi_6} \\ \delta e^{i\Delta\varphi_4} & \upsilon e^{i\Delta\varphi_5} & \zeta e^{i\Delta\varphi_6} & \alpha_4 \end{pmatrix} \qquad (54)$$

$$\begin{aligned} &\Delta\varphi_3 = \Delta\varphi_2 - \Delta\varphi_1 \\ &\Delta\varphi_5 = \Delta\varphi_4 - \Delta\varphi_1 \\ &\Delta\varphi_6 = \Delta\varphi_4 - \Delta\varphi_2 \end{aligned} \qquad (55)$$

Now we have

$$Tr\left(\hat{H}_4 \right) = -2, 0, 2 \qquad (56)$$

If $Tr\left(\hat{H}_4\right)=\pm 2$, the transition amplitudes (dispersions) are related to the spectrum through the following equations:

$$\begin{aligned}
\beta^2 &= \left(1\mp\alpha_1\right)\left(1\mp\alpha_2\right)\\
\gamma^2 &= \left(1\mp\alpha_1\right)\left(1\mp\alpha_3\right)\\
\delta^2 &= \left(1\mp\alpha_1\right)\left(1\mp\alpha_4\right)\\
\mu^2 &= \left(1\mp\alpha_2\right)\left(1\mp\alpha_3\right)\\
\upsilon^2 &= \left(1\mp\alpha_2\right)\left(1\mp\alpha_4\right)\\
\zeta^2 &= \left(1\mp\alpha_3\right)\left(1\mp\alpha_4\right)
\end{aligned} \tag{57}$$

Notice that these relations are universally valid and thus are subject to direct experimental verification.

As in the above, we may establish relations between eigenschaften and projection operators. For example, for $Tr\left(\hat{H}_4\right)=2$ we obtain

$$\begin{aligned}
\hat{I} &= \hat{P}_1+\hat{P}_2+\hat{P}_3+\hat{P}_4\\
\hat{H}_4^{(1)} &= -\hat{P}_1+\hat{P}_2+\hat{P}_3+\hat{P}_4=\hat{I}-2\hat{P}_1\\
\hat{H}_4^{(2)} &= \hat{P}_1-\hat{P}_2+\hat{P}_3+\hat{P}_4=\hat{I}-2\hat{P}_2\\
\hat{H}_4^{(3)} &= \hat{P}_1+\hat{P}_2-\hat{P}_3+\hat{P}_4=\hat{I}-2\hat{P}_3\\
\hat{H}_4^{(4)} &= \hat{P}_1+\hat{P}_2+\hat{P}_3-\hat{P}_4=\hat{I}-2\hat{P}_4
\end{aligned} \tag{58}$$

Again, we have

$$\frac{1}{2}\sum_{i=1}^{4}\hat{H}_4^{(i)}=\hat{I} \tag{59}$$

and

$$\begin{aligned}
&\hat{H}_4^{(1)}\cdot\hat{H}_4^{(2)}=\hat{H}_4^{(1)}+\hat{H}_4^{(2)}-\hat{H}_4^{(3)}-\hat{H}_4^{(4)}\\
&\left[\hat{H}_4^{(i)},\hat{H}_4^{(j)}\right]=0; i,j=1,2,3,4
\end{aligned} \tag{60}$$

and so on.

For the case $Tr\left(\hat{H}_4\right)=0$, we may write

$$\begin{aligned}
\hat{H}_4^{(1)} &= I\otimes\hat{H}_2\\
\hat{H}_4^{(2)} &= \hat{H}_2\otimes I\\
\hat{H}_4^{(3)} &= \hat{H}_2\otimes\hat{H}_2
\end{aligned} \tag{61}$$

since

$$Tr\left(A\otimes B\right)=TrA\cdot TrB \tag{62}$$

Then we have

$$\begin{aligned}
\hat{I} &= \hat{P}_1+\hat{P}_2+\hat{P}_3+\hat{P}_4\\
\hat{H}_4^{(1)} &= \hat{P}_1-\hat{P}_2+\hat{P}_3-\hat{P}_4\\
\hat{H}_4^{(2)} &= \hat{P}_1+\hat{P}_2-\hat{P}_3-\hat{P}_4\\
\hat{H}_4^{(3)} &= \hat{P}_1-\hat{P}_2-\hat{P}_3+\hat{P}_4
\end{aligned} \tag{63}$$

and

$$\begin{aligned}\hat{P}_1 &= \frac{1}{4}\left[\hat{I} + \hat{H}_4^{(1)} + \hat{H}_4^{(2)} + \hat{H}_4^{(3)}\right]\\ \hat{P}_2 &= \frac{1}{4}\left[\hat{I} - \hat{H}_4^{(1)} + \hat{H}_4^{(2)} - \hat{H}_4^{(3)}\right]\\ \hat{P}_3 &= \frac{1}{4}\left[\hat{I} + \hat{H}_4^{(1)} - \hat{H}_4^{(2)} - \hat{H}_4^{(3)}\right]\\ \hat{P}_4 &= \frac{1}{4}\left[\hat{I} - \hat{H}_4^{(1)} - \hat{H}_4^{(2)} + \hat{H}_4^{(3)}\right]\end{aligned} \quad (64)$$

Again we have

$$\begin{aligned}&\hat{H}_4^{(1)} \cdot \hat{H}_4^{(2)} = \hat{H}_4^{(3)}\\ &\left[\hat{H}_4^{(i)}, \hat{H}_4^{(j)}\right] = 0; i, j = 1,2,3\end{aligned} \quad (65)$$

We assume that the way to further generalization is obvious.

4. Holographic Detection: Quantum Reference Frames

Perhaps nobody needs an explanation of the mathematical formalism discussed in the previous section: we hope it speaks for itself. Nevertheless, we devote this section to the description of the physical "picture" behind the approach presented since that was the guideline that led us to it.

We address the following questions:

1) What is the difference between "on-off" and "or-and" switches in terms of quantum mechanical self-adjoint operators (observables)?

2) How are transition amplitudes between stationary (pure) states naturally and symmetrically incorporated within the amplitudes of these states?

3) Is it possible to measure $\bar{\hat{A}}$ and $\Delta\hat{A}$ simultaneously and how is the required setup arranged?

4) If it is possible, may measurements be performed using only macroscopic devices?

5) What does the Heisenberg Dispersion Relation (HDR) have to do with those measurements?

Our answer to the last question: almost nothing. It is well known [5] that the product of two noncommuting self-adjoint operators is not a self-adjoint operator and that the dispersion of their product is also not a self-adjoint operator. Therefore, there is no way to assign physical meaning to its numerical value. The theoretical importance of HDR tells us that quantum physics is the physics of extended objects and not the physics of material Newtonian points. The results of measurements are "pictures" and cannot in principle be treated as an image of a single point in space-time. The projection operators extensively used by J. von Neumann in his attempt to formulate his theory of measurements obviously play therole of "on-off" switches that define the basis of state vectors in Hilbert space. Therefore, it is reasonable to expect that "or-and" operators should be connected to them but in a slightly different way. Hadamard transformations [3] [6] which find their applications in image processing and quantum information theory seem to be suitable candidates. In addition, the notions of bits and qubitsnaturally appear as two-component wave packets. Finally, in order to provide a laboratory realization of the simultaneous measurement of the relevant amplitudes (relative generalized coordinates) and phase differences one should assure that wave packets arrive at every point on the detector screen.

Let us expose the content of our discussion to the eyes of the Schrödinger cat totally confused by the endless debates about its destiny. The usual justification for the apparent uncertainty refers to HDR. But empirical evidence tells us that the initial assumption that the cat may be considered as a quantum mechanical system containing inherent indeterminacy which then "becomes transformed into macroscopic indeterminacy" [7] is clearly wrong. If the state of the system (the "cat") is defined, one can measure its dispersion. Now, if in that given state the dispersion is not zero, we are dealing with an extended object and the expected result of the measurement should be represented by a picture of an unfortunate cat "mixed or smeared out in equal parts" [7]; if not, the cat was and will remain in the pure (definite) state, hopefully alive! Now, let us remember that in classical physics where only measurements of amplitudes are required, nobody doubts that the "moon is there" and that it is the same for all inertial reference frames, for example (**Figure 1**).

Here the lossless beam splitter is the macroscopic device which actively participates in the detection procedure ($\hat{H}_2^+ = \hat{H}_2$).

By contrast, in the microscopic quantum mechanical world (quantum optics) we are also required to measure the phase differences in order to obtain all existent and necessary information about the original object. This may be done using a similar setup, for example, see **Figure 2**.

However, in both cases the mirror and the lossless beam splitter participate only passively in the detection; they do not cause the wave function to collapse, but allow for extracting information on phase differences, since the referential component of the wave packet arrives a teach point of the detector screen

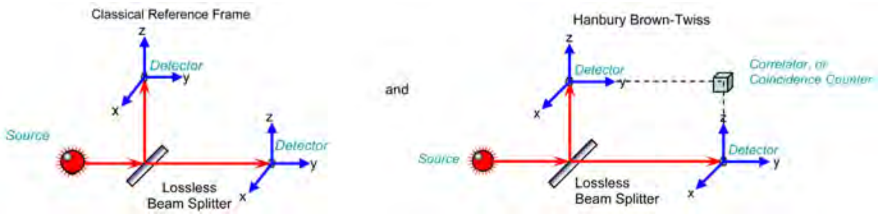


Figure 1. Classical optics measurement systems.

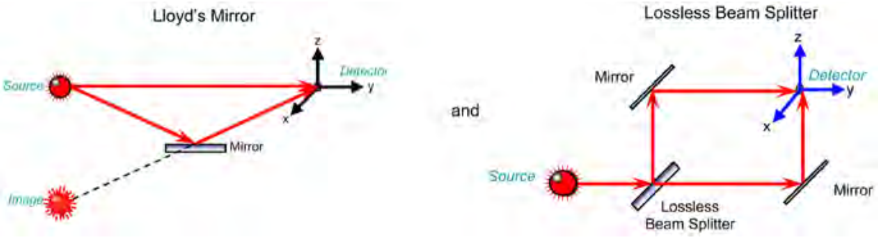


Figure 2. Quantum optics measurement systems.

together with the tested wave packet (within the inherent dispersion of the quantum mechanical space-time continuum). Then there is no reason to expect that the picture obtained would not provide an adequate image of the original object. It seems that now we are better equipped to formulate the dynamic (relativistic) laws of quantum physics. Ultimately that should lead to deeper understanding of the geometry of the space-time continuum.

Conflicts of Interest

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

References

[1] von Neumann, J. (1931) MathematischeGrundlagen der Quantenmechanik. Springer, Berlin.

[2] Dirac, P.A.M. (1958) The Principles of Quantum Mechanics. 4th Edition, Clarenton Press, Oxford.

[3] Aharonov, Y. and Vaidman, L. (1990) *Physical Review A*, **41**, 11. https://doi.org/10.1103/PhysRevA.41.11

[4] Courant, R. and Hilbert, D. (1931) Methoden der Mathematischen Physik. Springer, Berlin. https://doi.org/10.1007/978-3-642-47436-1

[5] Carruthers, P. and Nieto, M. (1968) *Reviews of Modern Physics*, **40**, 411. https://doi.org/10.1103/RevModPhys.40.411

[6] Bouwmeester, D., Ekker, A. and Zeilinger, A. (2000) The Physics of Quantum Information.

[7] Schrödinger, E. (1935) *Die Naturwissenschaften,* **48**, 807; **48**, 823; **48**, 844.

Journal of Modern Physics, 2018, 9, 2391-2401
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

New Discovery on Planck Units and Physical Dimension in Cosmic Continuum Theory

Xijia Wang

College of Mathematics and Statistics, Hunan Normal University, Changsha, China
Email: xijiawang@hunnu.edu.cn

How to cite this paper: Wang, X.J. (2018) New Discovery on Planck Units and Physical Dimension in Cosmic Continuum Theory. *Journal of Modern Physics*, **9**, 2391-2401.
https://doi.org/10.4236/jmp.2018.914153

Received: November 6, 2018
Accepted: December 4, 2018
Published: December 7, 2018

Abstract

In 1899, Max Planck integrated the Planck constant *h* with the gravitational constant *G* and the speed of light *c*, discovered a set of physical constants, and created Planck Units System. Since 20th century, the development of physics made the gravitational constant, the speed of light, and the Planck constant the most important fundamental constants of physics representing classical theory, relativity, and quantum theory, respectively. Now, the Planck Units have been given new physical meanings, revealing the mysteries of many physical boundaries. However, more than 100 years have passed, Planck Units System not only failed to get rid of the incompatibility between the basic theories of physics, but also cannot surpass the limitations of existing physics theories. In Cosmic Continuum Theory, physical dimensions can be transformed under the principle of equivalence. Planck units system not only integrates into the axiom system of Cosmic Continuum Theory, but also establishes a benchmark for the unity of physical dimensions. The introduction of the abstract physical dimensions "JX" and "XJ" makes the physical dimension of existence quantity and dimension quantity unified respectively.

Keywords

Cosmic Continuum, Axiomatic Physics, Fundamental Constants of Physics, Gravitational Redshift, Planck Units, Dimensional Analysis

1. Introduction

The fundamental constants of physics are the passwords of the universe and important scientific discoveries. Among all the physical constants, the well-known gravitational constant *G*, speed of light *c*, and Planck constant *h* are the three most fundamental constants of physics. These three fundamental constants of physics represent the different epoch-making basic theory of physics respectively. The gravitational constant G represents classical theory, the speed of light c represents relativity, and the Planck constant h represents quantum theory.

However, because these three basic theories of physics have different logical preconditions, they are regarded as three different cosmologies. There are disputes between high speed and low speed, macroscopic and microscopic, continuous and discrete between the three parties, and no consensus has been reached so far.

Cosmic Continuum Theory is an axiomatized physics system built on the mathematical continuum model. In Cosmic Continuum Theory, the universe is a continuum consisting of an existence continuum and an existing dimension continuum. The existence continuum is composed of mass bodies, energy bodies and dark mass bodies. The existing dimension continuum is composed of space, time and dark space. Mass, energy and dark mass are collectively called the existence quantity, and the quantity of space, time and dark space is called the dimension quantity [1] [2] [3].

This theory holds that a fundamental constant of physics should not only serve a certain physical basic theory. As cosmic codes of physics that gradually discovering in the course of continuous advancement, they are human's basic understanding of the natural world and should be an important basis for the unity of physics.

It was Max Planck himself who tried to unify fundamental constants of physics and achieved remarkable success. In 1899, he combined the Planck constant h with the gravitational constant G and the speed of light c, and found a series of physical constants. These physical constants are the so-called Planck units, and the following four constants are related to the existence quantity and the dimension quantity [4].

1) Planck mass: $m_p = \sqrt{hc/G} \approx 2.17651 \times 10^{-8}$ (kg).

2) Planck energy: $E_p = \sqrt{hc^5/G} \approx 1.9561 \times 10^{9}$ (J).

3) Planck time: $t_p = \sqrt{hG/c^5} \approx 5.39106 \times 10^{-44}$ (s).

4) Planck length: $l_p = \sqrt{hG/c^3} \approx 1.61619 \times 10^{-35}$ (m).

They are given corresponding physical meanings, such as: the mass of the ground state particles cannot be greater than the Plank mass; the energy cannot be greater than or equal to the Planck energy, otherwise it will collapse into a black hole; Planck time is an observable event Minimum process time; Planck length is a measure of Plank's quality black hole, unable to distinguish events within a distance less than Planck length; unable to describe events occurring within Planck time when the universe was born, etc. However, the Planck Unit cannot resolve the logical contradiction with the continuous space-time of the theory of relativity, because the space-time structure derived from Planck's length and Planck's time is discontinuous. Even for quantum field theory, Planck length and point particle models are in conflict [5]-[21].

2. The Physical Boundaries and Their Conversion of Planck Units System

In the following, we include the Planck units in the axiom system of Cosmic Continuum Theory for discussion.

2.1. Fundamental Concepts and Lemmas

This article uses the fundamental concepts and axioms system in [1]. The following lemmas are cited in [1], these lemmas will be used in later proofs.

Lemma 1: The existence quantity has its elementary unit $e_{\min}$ (Proof see [1]).

This lemma is equivalent to the quantification hypothesis.

Lemma 2: The interaction force between the initial fields of the existences Z_1 and Z_2 is: $f \sim Z_1 Z_2 / r^2$, and the action direction is along the line connecting Z_1 and Z_2, where, r is the distance between Z_1 and Z_2 (Proof see [1]).

This lemma contains the law of universal gravitation. When the existence is a mass body, the theorem is the law of universal gravitation.

Lemma 3: An existence has a maximum speed and the speed of an energy body is the maximum speed (Proof see [1]).

This lemma includes the speed of light postulate.

Lemma 4: An existence has a maximum frequency $\upsilon_{\max}$ (Proof see [1]).

Lemma 5: When the particle reaches its maximum speed, it will be converted to a quantum; when the particle or quantum reaches its maximum frequency, it will be transformed into dark particle (Proof see [1]).

Lemma 6: The existence quantities of mass m and energy $E = mc^2$ are equivalent: $m \equiv E$. "≡" is the "equivalent" symbol (Proof see [1]).

This lemmas contain mass-energy relation $E = mc^2$.

Lemma 7: The dimension quantity of 1-second of time is equivalent to that of c-kilometer space: 1 second ≡ c km, where, c indicates the speed of light (Proof see [1]).

Lemma 8: In a cosmic system A, the elementary particle, elementary quantum and elementary dark particle have equivalent inertia size: $m_{\min} \equiv q_{\min} \equiv d_{\min}$ (Axiom in [1]).

Lemma 9: A dimension quantity has its elementary units $w_{\min}$, and the dimension quantities of elementary space quantity $s_{\min}$, elementary time quantity $t_{\min}$ and elementary dark space quantity $w_{\min}$ are equivalent:
$w_{\min} \equiv s_{\min} \equiv t_{\min} \equiv g_{\min}$ (Proof see [1]).

Lemma 10: The existences are coupled with each other by energy, And in the presence of different structural levels, by the corresponding quantum from the role of convergence. If there is a structure at the structural level of the quantum of q connection, e the corresponding amount of existence, the corresponding structure of the particle m, dark particle d is also the amount of e: $m \equiv d \equiv q \equiv e$ (Axiom in [1]).

Lemma 11: There are only three basic forms of existence: particle, quantum, and dark particle (Axiom in [1]).

Lemma 12: The interaction force between the initial fields of the existences Z_1 and Z_2 is: $f \sim Z_1 Z_2 / r^2$, and the action direction is along the line connecting Z_1 and Z_2, where, *r* is the distance between Z_1 and Z_2 (Proof see [1]).

2.2. Derivation

1) Planck mass m_p

According to Lemma 2, the gravitational potential energy of two existing bodies with a distance of *r* and a mass of *m* is Gm^2/r, where *G* is the gravitational constant.

Let the vibration period of mass body *m* be *t*, then its frequency is:

$$\upsilon = 1/t \tag{1}$$

Also set its speed to *V*, then its movement distance in a period *t* is:

$$r = Vt \tag{2}$$

The search for Planck unit is to define the boundaries of physical events. According to Lemma 3, the velocity boundary of the microscopic particles is the speed of light *c*, so the change in the gravitational potential energy between the two bodies of Planck mass m_p in one cycle *t* should be:

$$\lim_{V \to c} Gm_p^2/(Vt) = Gm_p^2/(ct) \tag{3}$$

According to Lemma 1, the quantification hypothesis, we get:

$$Gm_p^2/(ct) = h\upsilon \tag{4}$$

Substituting (1) into (4) yields:

$$Gm_p^2/(ct) = h/t \tag{5}$$

From the formula (5), we can get the Planck mass m_p:

$$m_p = \sqrt{hc/G} \tag{6}$$

2) Planck energy E_p

Planck energy E_p can be obtained from Lemma 6:

$$E_p = m_p c^2 \tag{7}$$

Substituting (6) into (7)) gives Planck energy E_p:

$$E_p = \sqrt{hc^5/G} \tag{8}$$

3) Planck time t_p

According to Lemma 1, the Planck time t_p can be found by:

$$E_p = h\upsilon_p \tag{9}$$

where υ_p is the frequency of the Planck time t_p period:

$$\upsilon_p = 1/t_p \tag{10}$$

Substituting (10) into (9) yields:

$$E_p = h/t_p \tag{11}$$

Substituting (8) into (11) yields:

$$\sqrt{hc^5/G} = h/t_p \tag{12}$$

Obtain t_p from (12):

$$t_p = \sqrt{hG/c^5} \tag{13}$$

4) Planck length l_p

Planck length l_p can be obtained from Lemma 7:

$$l_p = t_p c \tag{14}$$

Substituting formula (13) into formula (14) yields:

$$l_p = \sqrt{hG/c^3} \tag{15}$$

2.3. Corollary

Conversion is one of the core ideas of Cosmic Continuum Theory. The mutual transformation of mass, energy, and dark mass makes the universe colorful. However, these transformations are not arbitrary, but are determined by the existence boundaries of particles, quantum, and dark particles. With the mutual conversion between particles, quantum, and dark particles, space, time, and dark space also follow.

Corollary 1: When the mass of particles reaches Planck mass m_p, the particles are converted into quantum and the mass is converted into energy.

Proof: From (3), we can see that Planck mass m_p is obtained by taking the speed as a variable to obtain the limit. According to Lemma 5, when the particle reaches the limit velocity *c*, it will turn into a quantum. According to the concept in [1], the energy body is a body composed of quantum, so when Planck mass is m_p, the mass has been transformed into energy. Q.E.D.

Corollary 2: When quantum energy reaches Planck energy E_p, quantum is transformed into dark particles and energy is converted into dark mass.

Proof: According to corollary 1, $m_p \equiv E_p$. From (3), we can see that Planck mass m_p is obtained by taking the particle velocity as a variable and finding the limit. However, it is impossible for the particles to reach the limit velocity *c*, so that the mass of the particles cannot reach Plank mass m_p, and thus the quantum cannot reach Planck energy E_p. According to Lemma 4, the quantum frequency has a maximum value. As a limit value, if the quantum energy reaches Planck energy E_p, the frequency υ_p of the quantum must reach a maximum value of $\upsilon_{\max}$. According to Lemma 5, when the quantum reaches the maximum frequency, it turns into a dark particle. According to the concept in [1], the dark mass is a body composed of dark particles. Therefore, when the quantum reaches a maximum frequency of $\upsilon_{\max}$, Planck energy E_p is converted to a dark mass. Q.E.D.

Corollary 3: Planck mass m_p is equivalent to Planck energy E_p: $m_p \equiv E_p$.

Proof: From (7), we can see that Planck energy E_p is obtained from Planck's mass m_p based on Lemma 6, according to Lemma 6, then $m_p \equiv E_p$. Q.E.D.

Corollary 4: When the amount of time is less than or equal to Planck time t_p, the time is converted to dark space.

Proof: From (9), we can see that Planck time t_p is derived from the Planck energy E_p and the frequency υ_p of the quantum. However, to reach Planck energy E_p, the frequency υ_p of the quantum must reach a maximum of $\upsilon_{\max}$. According to corollary 2, the quantum is transformed into dark particles and the energy is converted into dark mass. According to the concept in [1], dark space is the existence dimension of dark masses. Therefore, when the amount of time is less than or equal to Planck time t_p, the time will be converted into a dark space. Q.E.D.

Corollary 5: When the amount of space is less than or equal to Planck length l_p, the space is converted to a dark space.

Proof: From (14), Planck length l_p is the distance traveled at the speed of light in Planck time t_p. According to corollary 4, when time is less than or equal to Planck time t_p, time has been converted to dark space. Therefore, when the amount of space is less than or equal to Planck length l_p, the space will also be converted into a dark space. Q.E.D.

Corollary 6: Planck time t_p is equivalent to Planck length l_p: $t_p \equiv l_p$.

Proof: From (14), we can see that Planck length l_p is obtained from Planck time t_p based on Lemma 7. According to Lemma 7, then $t_p \equiv l_p$. Q.E.D.

The above 6 inferences fully demonstrate that Planck units do not have any logical contradictions and conflicts in the continuum of the universe and can perfectly express the physical boundaries and their transformation.

3. Unity of Physical Dimensions

The Planck unit system cleverly integrates the three fundamental constants of physics of Planck constant *h*, gravitational constant *G*, and speed of light *c*, and realized their dimensionless processing. Unfortunately, the physical significance of Planck units system is far from being fully reflected due to the inconsistency between the existing physical basic theories.

3.1. New Understanding of the Benchmark Value of Planck Units

The establishment of any physical dimension must be based on a certain benchmark. There are two kinds of benchmarks, one is the maximum, such as the speed of light *c*; the other is the minimum, such as the Planck constant *h*. There are many such benchmarks in the Planck units.

Corollary 7: The elementary unit of energy is $q_{\min} = h$ J. *h* is Planck constant.

Proof: According to Lemma 1 and Formula $E = h\upsilon$, where *E* refers to the energy of the quantum and υ refers to the frequency of the quantum, it can be known that Planck constant *h* is actually the energy of the unit frequency. This means that any energy body cannot be less than energy *h*, so Planck constant *h* is the elementary unit of energy. Q.E.D.

Corollary 8: The elementary unit of mass is $m_{\min} = h/c^2$ kg. *h* is Planck constant, and *c* is the speed of light.

According to Corollary 7, the elementary unit of energy is $q_{\min} = h$. According to Lemma 8, the inertia of elementary particles and elementary quanta in an universe system is equivalent: $m_{\min} \equiv q_{\min}$. According to Lemma 6, the existence quantities of mass *m* and energy $E = mc^2$ are equivalent: $m \equiv E$, so $h = m_{\min} c^2$. From this, we get: $m_{\min} = h/c^2$. Q.E.D.

Corollary 9: The elementary unit of time is $t_{\min} = t_p$ s. t_p is Planck time.

Proof: According to corollary 4, when the time is less than or equal to Planck time t_p, the time is converted to dark space. This shows that Planck time t_p is the minimum value of time. That is, the elementary unit of time dimension $t_{\min} = t_p$. Q.E.D.

Corollary 10: The elementary unit of space is $s_{\min} = l_p$ m. l_p is Planck length.

Proof: According to Corollary 5, when the length is less than or equal to Planck length l_p, the space is converted to dark space. This shows that Planck length l_p is the minimum value of space. That is, the elementary unit of length $s_{\min} = l_p$. Q.E.D.

Corollary 11: The maximum value of frequency is $\upsilon_{\max} = 1/t_p$ Hz. t_p is Planck time.

Proof: According to corollary 9, the elementary unit of time is $t_{\min} = t_p$, t_p is Planck time. This means that for any vibration cycle of the existing body *T*: $T \ge t_p$, and frequency $\upsilon = 1/T$, so for any frequency: $\upsilon = 1/T \le 1/t_p$, ie. $\upsilon_{\max} = 1/t_p$. Q.E.D.

Corollary 12: The maximum value of quantum energy is $E_{\max} = E_p$ J. E_p is Planck energy.

Proof: According to Corollary 11, the maximum value of frequency is $\upsilon_{\max} = 1/t_p$, t_p is the Planck time. The energy of quantum is $E = h\upsilon$, *h* is Planck constant, so $E_{\max} = h\upsilon_{\max} = h/t_p$. By (11), we can see that $E_p = h/t_p$, therefore $E_{\max} = E_p$. Q.E.D.

Corollary 13: The maximum value of particle mass is $m_{\max} = m_p$ kg. m_p is Planck mass.

Proof: According to Lemma 10, The existences are coupled with each other by energy, And in the presence of different structural levels, by the corresponding quantum from the role of convergence. If there is a structure at the structural level of the quantum of *q* connection, the corresponding structure of the particle *m* is also the amount of *q*: $m \equiv q$. According to corollary 3, Planck mass m_p is equivalent to Planck energy E_p: $m_p \equiv E_p$. This shows that m_p and E_p are in the same structure level. According to corollary 12, The maximum value of quantum energy is $E_{\max} = E_p$, E_p is Planck energy. That is, E_p is at the maximum structural level. Therefore, m_p is also at the structural level of maximum, that is, m_p is the maximum value of particle mass: $m_{\max} = m_p$. Q.E.D.

Corollary 7-13 allows us to obtain the elementary unit of energy, the elementary unit of mass, the elementary unit of time, the elementary unit of space, the maximum of frequency, the maximum of quantum energy, and the maximum of particle mass 7 limit constants, this shows that the Cosmic Continuum Theory further develops the benchmark value of the Planck units.

3.2. Equivalent Abstract Physical Dimension

In the system of Planck units, the values of *h*, *G*, *c* and Planck units are all equal to 1, and the mass-energy equation is simplified to $E = m$. It looks a bit similar to Lemma 6: the existence quantities of mass *m* and energy $E = mc^2$ are equivalent: $m \equiv E$. In reality, it is not. The former is a numerical simplification and is a non-dimensionalized process, represented by the "=" symbol. The latter is the physical equivalent of mass and energy, and it is the unity of physical dimensions, represented by the "≡" symbol. In Lemma 6, the physical dimension of *m* is kilogram, and the physical dimension of *E* is Joule; When it satisfies $E = mc^2$, $m \equiv E$ expresses: $m\,\text{kg} \equiv E\,\text{J}$.

Corollary 14: The physical dimensions of existence quantity can be unified on abstract physical dimensions.

Proof: According to Lemma 1, Lemma 8, the elementary unit of mass $m_{\min}$, the elementary unit of energy $q_{\min}$, and the elementary unit of dark quality $d_{\min}$ equivalence, which results in the existence of the elementary unit of quantity $e_{\min}$: $e_{\min} \equiv m_{\min} \equiv q_{\min} \equiv d_{\min}$.

Set $e_{\min} = 1\,\text{JX}$, then:

$$1\,\text{JX} \equiv m_{\min} = h/c^2\,(\text{kg}) \tag{16}$$

$$1\,\text{JX} \equiv q_{\min} = h\,(\text{J}) \tag{17}$$

$$1\,\text{JX} \equiv d_{\min}\ \text{Unit} \tag{18}$$

Due to the fact that there is no international standard unit for the amount of dark mass, it is temporarily replaced by "Unit". According to Lemma 11, there are only particle, quantum, and dark particle three basic forms, thus we have unified the physical dimension of existence quantity to an abstract physical dimension "JX":

$$e_{\min} = 1\,\text{JX} \equiv h/c^2\ \text{kg} \equiv h\ \text{J} \equiv d_{\min}\ \text{Unit} \tag{19}$$

So the physical dimensions of existence quantity can be unified on abstract physical dimensions. Q.E.D.

Corollary 15: The physical dimensions of dimension quantity can be unified on abstract physical dimension.

Proof: According to Lemma 9, the elementary unit of space $s_{\min}$, the elementary unit of time $t_{\min}$, and the elementary unit of dark space $g_{\min}$ are equivalent, so as to produce the elementary unit of dimension $w_{\min}$: $w_{\min} \equiv s_{\min} \equiv t_{\min} \equiv g_{\min}$.

Set $w_{\min} = 1\,\text{XJ}$, then:

$$1\ \mathrm{XJ} \equiv t_{\min} = t_p = \sqrt{hG/c^5}\ (\mathrm{s}) \tag{20}$$

$$1\ \mathrm{XJ} \equiv s_{\min} = l_p = \sqrt{hG/c^3}\ (\mathrm{m}) \tag{21}$$

$$1\ \mathrm{XJ} \equiv g_{\min}\ \mathrm{Unit} \tag{22}$$

Similarly, there is currently no international standard unit for the amount of dark space, it is temporarily replaced by "Unit". According to Lemma 11, there are only three basic forms of particle, quantum, and dark particle. According to Concept 4, Concept 5, and Concept 6, there are also only three basic dimensional forms: space, time, and dark space. Therefore, we have unified the physical dimension of dimension quantity to an abstract physical dimension "XJ":

$$w_{\min} = 1\ \mathrm{XJ} \equiv \sqrt{hG/c^5}\ \mathrm{s} \equiv \sqrt{hG/c^3}\ \mathrm{m} \equiv d_{\min}\ \mathrm{Unit} \tag{23}$$

So the physical dimensions of dimension quantity can be unified on abstract physical dimension. Q.E.D.

Corollary 16: All changes in the universe can be equivalent on abstract physical dimensions.

Proof: All changes in the universe are changes in the quantity and dimension of existence. According to Corollary 14, the physical dimensions of existence quantity can be unified on abstract physical dimensions; and according to Corollary 15, the physical dimensions of dimension quantity can be unified on abstract physical dimension. According to Lemma 8 and Lemma 9, the unity of physical dimensions is based on physical equivalence, so all changes in the universe can be equivalent on abstract physical dimensions. Q.E.D.

Obviously, this physical dimensions system is exactly the same as the Planck Units System. However, the two physical dimensions "JX" and "XJ" are essentially different from our existing physical dimensions. Existing physical dimensions are specific physical dimensions, including Planck units. There is no connection between different specific physical dimensions. But the "JX" and "XJ" are abstract physical dimensions. The function of an abstract physical dimension is to achieve the unity of specific physical dimensions.

3.3. A New Interpretation of Gravitational Redshift

The following is an attempt to solve the gravitational redshift problem by applying the equivalent unified scheme of physical dimensions proposed in this paper.

We know that the energy of a photon E is:

$$E = h\upsilon \tag{24}$$

h is the Planck constant and υ is the photon frequency. According to Lemma 6, The existence quantities of mass m and energy $E = mc^2$ are equivalent: $m \equiv E$. Therefore:

$$E = h\upsilon \equiv mc^2 \tag{25}$$

c is the speed of light. So we can get:

$$m \equiv h\upsilon/c^2 \tag{26}$$

According to Lemma 12, it can be seen that the photon escapes from a certain celestial body M to a distant place of R, and the change of gravitational potential energy E_p is:

$$E_p = -GMm/R \tag{27}$$

G is the gravitational constant. According to corollary 14, the physical dimensions of existence quantity can be unified on abstract physical dimensions. Substituting (26) into (27) gives:

$$E_p = -GMh\upsilon/Rc^2 \tag{28}$$

The amount of change in the gravitational potential energy of a photon is the amount of energy change of this photon. According to the formula (24), if the photon frequency change amount is υ_p, then:

$$\upsilon_p = E_p/h \tag{29}$$

Substituting (28) into (29):

$$\upsilon_p = -GM\upsilon/Rc^2 \tag{30}$$

Deform (30) to get:

$$\upsilon_p/\upsilon = -GM/Rc^2 \tag{31}$$

The result is consistent with the gravitational redshift formula derived by general relativity. This shows that the introduction of the concept of relativistic mass is equivalent to the unification of physical dimensions.

4. Conclusion

The new fundamental constants of physics often represent a new basic theory of physics, and the logical unity of basic physical constants is essentially the unity of physics. The Planck system of units profoundly reveals the physical boundary problem, but it has always been constrained by inconsistencies between the basic theories of physics. In Cosmic Continuum Theory, Planck Units System completely escapes the incompatibility of the physical basis. The axiomatization of Planck Units System enables the logical unity of fundamental constants of physics that represent different physical foundations to be realized. In particular, based on Planck unit, 7 limit constants were deduced, giving a unified benchmark for physical dimensions. The introduction of abstract physical dimensions "JX" and "XJ" to achieve unity of physical dimensions will sweep away another barrier for the unification of physics.

Conflicts of Interest

I declare that there is no competing Interest.

References

[1] Wang, X.J. (2018) *Journal of Modern Physics*, **9**, 1250-1270. https://doi.org/10.4236/jmp.2018.96074

[2] Wang, X.J. and Wu, J.X. (1992) The Unity Theory. Haitian Publishing House, Shenzhen.

[3] Wang, X.J. and Wu, J.X. (2001) Crack to the Puzzle of Scientific Unity. Hunan Science & Technology Press, Changsha.

[4] Planck, M. (1899) Über irreversible Strahlungsvorgänge. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin, 5: 440-480. (pp. 478-480 contain the first appearance of the Planck base units other than the Planck charge, and of Planck's constant, which Planck denoted by b)

[5] Wesson, P.S. (1980) *Space Science Reviews*, **27**, 109-153.

[6] Duff, M., Okun, L.B. and Veneziano, G. (2002) *Journal of High Energy Physics*, **3**, 23. https://doi.org/10.1088/1126-6708/2002/03/023

[7] Ashtekar, A. (2013) Loop Quantum Gravity and the The Planck Regime of Cosmology, arXiv:1303.4989.

[8] Gonzalez-Mestres, L. (2014) *EPJ Web of Conferences*, **71**, 00063.

[9] Post, E.J. (1982) *Foundations of Physics*, **12**, 169-195. https://doi.org/10.1007/BF00736847

[10] Robinett, R.W. (2015) *American Journal of Physics*, **83**, 353-361. https://doi.org/10.1119/1.4902882

[11] Fischer, J. and Ullrich, J. (2016) *Nature Physics*, **12**, 4-7. https://doi.org/10.1038/nphys3612

[12] Kisak, P.F. (2015) Planck Units: The Fundamental Scale of Cosmology. CreateSpace Independent Pub, Colorado Springs.

[13] CODATA Value: Newtonian Constant of Gravitation. The NIST Reference on Constants, Units, and Uncertainty. US National Institute of Standards and Technology. June 2015. Retrieved 2015-09-25. 2014 CODATA Recommended Values.

[14] Staff. Birth of the Universe. University of Oregon. Retrieved September 24, 2016. (Discusses "Planck time" and "Planck era" at the very beginning of the Universe)

[15] Barrow, J.D. (2002) The Constants of Nature; From Alpha to Omega—The Numbers that Encode the Deepest Secrets of the Universe. Pantheon Books, New York.

[16] Barrow, J.D. and Tipler, F.J. (1988) The Anthropic Cosmological Principle. Oxford University Press, Oxford.

[17] Dirac, P.A.M. (1938) *Proceedings of the Royal Society A*, **165**, 199-208. https://doi.org/10.1098/rspa.1938.0053

[18] Tomilin, K.A. (1999) Natural Systems of Units: To the Centenary Anniversary of the Planck System, 287-296.

[19] Pavšic, M. (2001) The Landscape of Theoretical Physics: A Global View. Kluwer Academic, Dordrecht, 347-352. https://arxiv.org/abs/gr-qc/0610061

[20] Davies, P.C., Davis, T.M. and Lineweaver, C.H. (2002) *Nature*, **418**, 602-603.

[21] Nieto, J., Poupaud, F. and Soler, J. (2001) *Archive for Rational Mechanics and Analysis*, **158**, 29-59. https://doi.org/10.1007/s002050100139

Journal of Modern Physics, 2018, 9, 2402-2419
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Fluid State of Dirac Quantum Particles

Vu B. Ho

Advanced Study, 9 Adela Court, Mulgrave, Victoria, Australia
Email: vubho@bigpond.net.au

How to cite this paper: Ho, V.B. (2018) Fluid State of Dirac Quantum Particles. *Journal of Modern Physics*, **9**, 2402-2419.
https://doi.org/10.4236/jmp.2018.914154

Received: November 7, 2018
Accepted: December 10, 2018
Published: December 13, 2018

Abstract

In our previous works, we suggest that quantum particles are composite physical objects endowed with the geometric and topological structures of their corresponding differentiable manifolds that would allow them to imitate and adapt to physical environments. In this work, we show that Dirac equation in fact describes quantum particles as composite structures that are in a fluid state in which the components of the wavefunction can be identified with the stream function and the velocity potential of a potential flow formulated in the theory of classical fluids. We also show that Dirac quantum particles can manifest as standing waves which are the result of the superposition of two fluid flows moving in opposite directions. However, for a steady motion a Dirac quantum particle does not exhibit a wave motion even though it has the potential to establish a wave within its physical structure, therefore, without an external disturbance a Dirac quantum particle may be considered as a classical particle defined in classical physics. And furthermore, from the fact that there are two identical fluid flows in opposite directions within their physical structures, the fluid state model of Dirac quantum particles can be used to explain why fermions are spin-half particles.

Keywords

Dirac Equation, Wave Mechanics, Stan, Fluid Mechanics, Stream Function, Velocity Potential, Potential Flow, General Relativity, Maxwell Field Equations, CW Complexes, Differential Geometry, Topology, Differentiable Manifolds, Topological Transformation

1. Introductory Summary

In our previous works on spacetime structures of quantum particles, we suggest that quantum particles should be endowed with geometric and topological structures of differentiable manifolds and their motion should be described as isometric embeddings in higher Euclidean space. We also suggest that all quantum particles are formed from mass points which are joined together by contact forces as a consequence of viewing quantum particles as CW-complexes [1] [2] [3]. Fundamentally, we show that the three main dynamical descriptions of physical events in classical physics, namely Newton mechanics, Maxwell electromagnetism and Einstein gravitation, can be formulated in the same general covariant form and they can be represented by the general equation

$$\nabla_\beta M = kJ \tag{1}$$

where M is a mathematical object that represents the corresponding physical system and ∇_β is a covariant derivative. For Newton mechanics, $M = \frac{1}{2} m \sum_{\mu=1}^{3} \left(\mathrm{d}x^\mu / \mathrm{d}t \right)^2 + V$ and $J = 0$. For Maxwell electromagnetism, $M = F^{\alpha\beta} = \partial^\mu A^\nu - \partial^\nu A^\mu$, with the four-vector potential $A^\mu \equiv (V, \mathbf{A})$ and J can be identified with the electric and magnetic currents. And for Einstein gravitation, $M = R^{\alpha\beta}$ and J can be defined in terms of a metric $g_{\alpha\beta}$ and the Ricci scalar curvature using the Bianchi identities $\nabla_\beta R^{\alpha\beta} = \frac{1}{2} g^{\alpha\beta} \nabla_\beta R$, that is, $J = \frac{1}{2} g^{\alpha\beta} \nabla_\beta R$. If we use the Bianchi identities as field equations for the gravitational field then Einstein field equations $T_{\mu\nu} = k \left(R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} \right)$, as in the case of the electromagnetic field, should be regarded as a definition for the energy-momentum tensor $T_{\mu\nu}$ for the gravitational field [4]. An interesting feature that emerges from Equation (1) for the gravitational field is that we can derive the Ricci flow $\partial g_{\alpha\beta} / \partial t = \kappa R_{\alpha\beta}$ for a vacuum field $J = 0$. Mathematically, the Ricci flow is a geometric process that can be employed to smooth out irregularities of a Riemannian manifold [5]. From the definition of the four-current $j^\alpha = (\rho, \mathbf{j}_i) = \frac{1}{2} g^{\alpha\beta} \nabla_\beta R$ for the gravitational field, by comparing with the Poisson equation for a potential V in classical physics, $\nabla^2 V = 4\pi\rho$, we can identify the scalar potential V with the Ricci scalar curvature R and then obtain a diffusion equation $\partial_t R = k \nabla^2 R$ whose solutions can be found to take the form $R(x, y, z, t) = \left(M / \left(\sqrt{4\pi k t} \right)^3 \right) \mathrm{e}^{-\left(x^2 + y^2 + z^2 \right) / 4kt}$, which determines the probabilistic distribution of an amount of geometrical substance M which is defined via the Ricci scalar curvature R and manifests as observable matter [6]. It is worth mentioning that in fact a similar diffusion equation can also be derived from the Ricci flow $\partial g_{\alpha\beta} / \partial t = \kappa R_{\alpha\beta}$ of the form $\partial R / \partial t = \Delta R + 2 |\mathrm{Ric}|^2$, where Δ is the Laplacian defined as $\Delta = g^{\alpha\beta} \nabla_\alpha \nabla_\beta$ and $|\mathrm{Ric}|$ is a shorthand for a mathematical expression that we will not be concerned with in this work [7]. Therefore, the Bianchi field equations of general relativity in the covariant form given in Equation (1) can be used to formulate quantum particles as differentiable manifolds. For example, we showed that the Ricci scalar curvature R associated with a differentiable manifold that represents a quantum system, such as the hydrogen atom, can be expressed in terms of the Schrödinger wavefunction ψ in quantum mechanics as $R = k \left(\sum_{\mu=1}^{3} \left(\mathrm{d}x^\mu / \mathrm{d}t \right)^2 - (\hbar / m) \left(\partial_t \psi + \sum_{\mu=1}^{3} \partial_\mu \psi \left(\mathrm{d}x^\mu / \mathrm{d}t \right) / \psi \right) \right)$.

On the other hand, we have also shown that Maxwell field equations of electromagnetism and Dirac relativistic equation of quantum mechanics can be formulated covariantly from a general system of linear first order partial differential equations [8] [9]. An explicit form of a system of linear first order partial differential equations can be written as follows [10] [11]

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^{r}\frac{\partial\psi_i}{\partial x_j}=k_1\sum_{l=1}^{n}b_l^r\psi_l+k_2c^r, r=1,2,\cdots,n \tag{2}$$

The system of equations given in Equation (2) can be rewritten in a matrix form as

$$\left(\sum_{i=1}^{n}A_i\frac{\partial}{\partial x_i}\right)\psi=k_1\sigma\psi+k_2J \tag{3}$$

where $\psi=(\psi_1,\psi_2,\cdots,\psi_n)^{\mathrm{T}}$, $\partial\psi/\partial x_i=(\partial\psi_1/\partial x_i,\partial\psi_2/\partial x_i,\cdots,\partial\psi_n/\partial x_i)^{\mathrm{T}}$, A_i, σ and J are matrices representing the quantities a_{ij}^k, b_l^r and c^r, and k_1 and k_2 are undetermined constants. Now, if we apply the operator $\sum_{i=1}^{n}A_i\partial/\partial x_i$ on the left on both sides of Equation (3) then we obtain

$$\begin{aligned}&\left(\sum_{i=1}^{n}A_i^2\frac{\partial^2}{\partial x_i^2}+\sum_{i=1}^{n}\sum_{j>i}^{n}\left(A_iA_j+A_jA_i\right)\frac{\partial^2}{\partial x_i\partial x_j}\right)\psi\\&=k_1^2\sigma^2\psi+k_1k_2\sigma J+k_2\sum_{i=1}^{n}A_i\frac{\partial J}{\partial x_i}\end{aligned} \tag{4}$$

In order for the above systems of partial differential equations to be used to describe physical phenomena, the matrices A_i must be determined. We have shown that for both Dirac and Maxwell field equations, the matrices A_i must take a form so that Equation (4) reduces to the following equation

$$\left(\sum_{i=1}^{n}A_i^2\frac{\partial^2}{\partial x_i^2}\right)\psi=k_1^2\sigma^2\psi+k_1k_2\sigma J+k_2\sum_{i=1}^{n}A_i\frac{\partial J}{\partial x_i} \tag{5}$$

To obtain Dirac equation we simply set $A_iA_j+A_jA_i=0$ with $A_i^2=\pm1$, and in this case the matrices A_i are the matrices γ_i given as [12]

$$\begin{aligned}&\gamma_1=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix},\gamma_2=\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&-1&0&0\\-1&0&0&0\end{pmatrix},\\&\gamma_3=\begin{pmatrix}0&0&0&-i\\0&0&i&0\\0&i&0&0\\-i&0&0&0\end{pmatrix},\gamma_4=\begin{pmatrix}0&0&1&0\\0&0&0&-1\\-1&0&0&0\\0&1&0&0\end{pmatrix}\end{aligned} \tag{6}$$

For Maxwell field equations, in order to specify the matrices A_i we need to use the form of Maxwell field equations established in classical electrodynamics [13] [14]. And the matrices A_i take the forms

$$A_1=\begin{pmatrix}-1&0&0&0&0&0\\0&-1&0&0&0&0\\0&0&-1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{pmatrix}, A_2=\begin{pmatrix}0&0&0&0&0&0\\0&0&0&0&0&-1\\0&0&0&0&1&0\\0&0&0&0&0&0\\0&0&-1&0&0&0\\0&1&0&0&0&0\end{pmatrix},$$
$$A_3=\begin{pmatrix}0&0&0&0&0&1\\0&0&0&0&0&0\\0&0&0&-1&0&0\\0&0&1&0&0&0\\0&0&0&0&0&0\\-1&0&0&0&0&0\end{pmatrix}, A_4=\begin{pmatrix}0&0&0&0&-1&0\\0&0&0&1&0&0\\0&0&0&0&0&0\\0&-1&0&0&0&0\\1&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix}, \tag{7}$$
$$A_5=\begin{pmatrix}\mu&0&0&0&0&0\\0&\mu&0&0&0&0\\0&0&\mu&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{pmatrix}$$

Besides the covariant formulations of classical and quantum physics as described above, we have also discussed the topological transformation of quantum dynamics by showing the wave dynamics of a quantum particle on different types of topological structures in various dimensions from the fundamental polygons of the corresponding universal covering spaces [15]. We presented our discussions in the form of Bohr model in one, two and three dimensions using linear wave equations. For the clarity of our presentation in terms of Bohr model, we want to mention here that in order to successfully construct a model for the hydrogen atom which predicts correctly the spectrum of the energy radiated from the atom, Bohr proposed three postulates which state that the centripetal force required for the electron to orbit the nucleus in a stable circle is the Coulomb force $mv^2/r = kq^2/r^2$, the permissible orbits are those that satisfy the condition that the angular momentum of the electron equals $n\hbar$, that is $mvr = n\hbar$, and when the electron moves in one of the stable orbits it does not radiate, however, it will radiate when it makes a transition between the stable orbits [16]. Furthermore, in his work on the concept of matter wave, de Broglie proposed that an electron has both a wave and a particle nature by regarding the electron as a standing wave around the circumference of an orbit, as shown in **Figure 1** [17].

It is seen that de Broglie's requirement leads to the wave condition $2\pi r = n\lambda$. This is equivalent to assuming that the standing wave around a circle, which is a 1-sphere, is similar to a standing wave on the fundamental interval of a straight line which is the universal covering space of the circle S^1, where the translations taking the interval to the next images will generate the holonomy group [18]. If we apply de Broglie wavelength λ defined in terms of the momentum

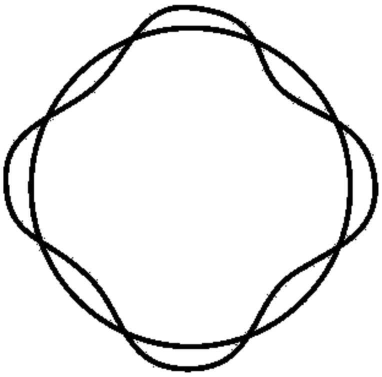


Figure 1. de Broglie waves around a circle.

of a quantum particle $p = mv$ as $\lambda = h/mv$, then using the wavelength also given by $\lambda = 2\pi R/n$ we obtain $h/mv = 2\pi R/n$, that leads to the Bohr's postulate of the quantisation of angular momentum $mvR = n\hbar$. Then the energy spectrum E_n can be calculated from Coulomb's law $mv^2/R = kq^2/R^2$ as $E_n = -mk^2q^4/2\hbar^2n^2$, where R now is the radius of the nth stationary orbit.

In mathematics, the fundamental polygon in one dimension is an interval and the universal covering space is the straight line and in this case the standing wave on a finite string is transformed into the standing wave on a circle which can be applied into the Bohr model of the hydrogen atom. In two dimensions, the fundamental polygon is a square and the universal covering space is the plane and in this case the standing wave on the square is transformed into the standing wave on different surfaces that can be formed by gluing opposite sides of the square, which include a 2-sphere, a 2-torus, a Klein bottle and a projective plane. This may be seen as an extension of the Bohr model of the hydrogen atom from one-dimensional manifolds of the 1-sphere and 1-torus embedded in the ambient two-dimensional Euclidean space R^2 into two-dimensional manifolds embedded or immersed in the ambient three-dimensional Euclidean space R^3. In three dimensions, the fundamental polygon is a cube and the universal covering space is the three-dimensional Euclidean space. It is shown that a 3-torus and the manifold $K \times S^1$ defined as the product of a Klein bottle and a circle can be constructed by gluing opposite faces of a cube therefore in three-dimensions the standing wave on a cube is transformed into the standing wave on a 3-torus or on the manifold $K \times S^1$. We also discuss a transformation of a stationary wave on the fundamental cube into a stationary wave on a 3-sphere despite it still remains unknown whether a 3-sphere can be constructed directly from a cube by gluing its opposite faces. In spite of this uncertainty, however, we speculate that mathematical degeneracy in which an element of a class of objects degenerates into an element of a different but simpler class may play an important role in quantum dynamics. For example, a 2-sphere is a degenerate 2-torus when the axis of revolution passes through the centre of the generating circle. Therefore, it seems reasonable to assume that if an n-torus degenerates into an n-sphere then wavefunctions on an n-torus may also be degenerated into wavefunctions on an

n-sphere. Furthermore, since an n-sphere can degenerate itself into a single point, therefore the mathematical degeneracy may be related to the concept of wavefunction collapse in quantum mechanics where the classical observables such as position and momentum can only be obtained from the collapse of the associated wavefunctions for physical measurements. This consideration suggests that quantum particles associated with differentiable manifolds may possess the more stable mathematical structures of an n-torus rather than those of an n-sphere.

The above formulation of quantum particles in terms of differentiable manifolds and the consideration of their intrinsic geometric and topological characteristics raise the question of how the standing waves that represent quantum particles could be established physically. The aim of this work is to answer this question by showing that Dirac equation in fact describes quantum particles as composite structures that are in a fluid state in which the components of the wavefunction can be identified with the stream function and the velocity potential of a potential flow formulated in the theory of classical fluids. In this case Dirac quantum particles can manifest as standing waves which are the result of the superposition of two fluid flows moving in opposite directions. For example, if two opposite waves are represented by the function $\psi_1(r,t) = a\sin(kr - vt)$ and $\psi_2(r,t) = a\sin(kr + vt)$ then the resultant standing wave can be obtained as $\psi_1(r,t) + \psi_2(r,t) = 2a\sin(kr)\cos(vt)$. We also show that even though a Dirac quantum particle has the potential to transfer energy in opposite directions to establish a standing wave within its physical structure, for a steady motion without an external disturbance a Dirac quantum particle can be considered as a classical particle defined in classical physics.

2. Two-Dimensional Hydrogen-Like Physical System

In this work we will show that quantum particles can be described as physical systems in a state of fluids that can be formulated in terms of the fluid dynamics in two dimensions. Since such systems can be seen to have the physical structure of a two-dimensional hydrogen-like atom therefore in this section we will examine further how they can be described in terms of quantum mechanics. First we need to extend our formulation of Maxwell field equations of electromagnetism and Dirac relativistic equation from a general system of linear first order partial differential equations to that with an external field. Such system of equations is given as follows [19]

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^{r}\frac{\partial \psi_i}{\partial x_j} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} b_{ij}^{r}V_j + c_i^{r}\right)\psi_i + d^{r},\ r = 1, 2, \cdots, n \tag{8}$$

The system of equations given in Equation (8) can be rewritten in a matrix form as

$$\left(\sum_{i=1}^{n} A_i \frac{\partial}{\partial x_i}\right)\psi = -i\left(\sum_{i=1}^{n} qB_iV_i + m\sigma\right)\psi + J \tag{9}$$

where $\psi=\left(\psi_1,\psi_2,\cdots,\psi_n\right)^{\mathrm{T}}$, $\partial\psi/\partial x_i=\left(\partial\psi_1/\partial x_i,\partial\psi_2/\partial x_i,\cdots,\partial\psi_n/\partial x_i\right)^{\mathrm{T}}$ with A_i, B_i, σ and J are matrices representing the quantities a_{ij}^r, b_{ij}^r, c_j^r and d^r, which are assumed to be constant in this work. While the quantities q, m and J represent physical entities related directly to the physical properties of the particle, and the quantities V_i represent an external field, such as the potentials of an electromagnetic field. By applying the operator $\sum_{i=1}^n A_i\partial/\partial x_i$ on the left on both sides of Equation (9), with the assumption that the coefficients a_{ij}^k, b_i^r and c^r are constants, then we obtain

$$\left(\sum_{i=1}^n A_i\frac{\partial}{\partial x_i}\right)\left(\sum_{j=1}^n A_j\frac{\partial}{\partial x_j}\right)\psi=\left(\sum_{i=1}^n A_i\frac{\partial}{\partial x_i}\right)\left(-i\left(\sum_{j=1}^n qB_jV_j+m\sigma\right)\psi+J\right)\tag{10}$$

Since the quantities A_i, B_i, σ, q, m and J are assumed to be constant, Equation (10) becomes

$$\begin{aligned}&\left(\sum_{i=1}^n A_i^2\frac{\partial^2}{\partial x_i^2}+\sum_{i=1}^n\sum_{j>i}^n\left(A_iA_j+A_jA_i\right)\frac{\partial^2}{\partial x_i\partial x_j}\right)\psi\\&=\left(-i\left(\sum_{i=1}^n A_i\frac{\partial}{\partial x_i}\right)\left(\sum_{j=1}^n qB_jV_j+m\sigma\right)\right)\psi\\&\quad-i\left(\sum_{i=1}^n qB_iV_i+m\sigma\right)\left(\left(\sum_{j=1}^n A_j\frac{\partial}{\partial x_j}\right)\psi\right)+\sum_{i=1}^n A_i\frac{\partial J}{\partial x_i}\\&=-i\left(\sum_{i=1}^n\sum_{j=1}^n qA_iB_j\frac{\partial V_j}{\partial x_i}\right)\psi-\left(\sum_{i=1}^n\sum_{j>i}^n q^2\left(B_iB_j+B_jB_i\right)V_iV_j\right.\\&\quad\left.-2i\sum_{i=1}^n qmB_iV_i\sigma-m^2\sigma^2\right)\psi-i\left(\sum_{i=1}^n qB_iV_i+m\sigma\right)J+\sum_{i=1}^n A_i\frac{\partial J}{\partial x_i}\end{aligned}\tag{11}$$

Dirac equation for an arbitrary field can be formulated from the system of linear first order partial differential equations given in Equation (9) by setting $B_i=A_i=\gamma_i$, $\sigma=1$, $J=0$ and $A_iA_j+A_jA_i=0$. In this case, in terms of the operator γ_i, Equation (9) becomes

$$\left(\sum_{i=1}^4\gamma_i\frac{\partial}{\partial x_i}\right)\psi=-i\left(\sum_{i=1}^4 q\gamma_iV_i+m\right)\psi\tag{12}$$

Equation (12) can be written in a covariant form as Dirac equation for an arbitrary field as

$$\left(\gamma^\mu\left(i\partial_\mu-qV_\mu\right)-m\right)\psi=0\tag{13}$$

Equation (11) also reduces to the following equation

$$\left(\sum_{i=1}^4\gamma_i^2\frac{\partial^2}{\partial x_i^2}\right)\psi=\left(-i\sum_{i=1}^4\sum_{j>i}^4 q\gamma_i\gamma_j\left(\frac{\partial V_j}{\partial x_i}-\frac{\partial V_i}{\partial x_j}\right)+2i\sum_{i=1}^4 qm\gamma_iV_i-m^2\right)\psi\tag{14}$$

Even though in the following we will examine only physical states of Dirac quantum particles in which $V_i=0$ where the physical quantity V_i is assumed to be associated with an external field, however, if we consider quantum particles as differentiable manifolds which are formed by mass points joined together by contact forces then we may suggest that they are endowed with intrinsic geometric and topological structures and in this case the quantity V_i may be considered as an internal field that is responsible for the stability of the physical structure of a quantum particle. As we will show below this is in fact the case when at least part of a quantum particle exists as a two-dimensional structure in which the intrinsic angular momentum can take half-integer values. The problem that we considered can actually be started with Dirac equation given in Equation (13). It can be shown that in the non-relativistic limit, Dirac equation reduces to the Pauli equation for stationary system as [20]

$$\left(\frac{1}{2\mu}\left(-i\hbar\nabla - q\mathbf{A}\right)^2 - \frac{q\hbar}{2\mu}(\boldsymbol{\sigma}\cdot\mathbf{B}) + q\phi\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{15}$$

In the case when $\mathbf{A} = 0$, $\mathbf{B} = 0$, and $q\phi = k/r$ then we have

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi(\mathbf{r}) - \frac{k}{r}\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{16}$$

Now let us examine a physical system that is described by the Schrödinger wave equation given in Equation (16) from the viewpoint of an observer who sees it as a planar system [21] [22] [23]. It is shown that if we consider physical systems whose configuration space is multiply connected, such as the physical system of a hydrogen-like atom in two rather than three dimensions, then multivalued wavefunctions can be used to describe the system [24]. In two-dimensional space, the Schrödinger equation in the planar polar coordinates takes the form

$$-\frac{\hbar^2}{2\mu}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2}\right]\psi(r,\phi) - \frac{k}{r}\psi(r,\phi) = E\psi(r,\phi) \tag{17}$$

Solutions of the form $\psi(r,\phi) = R(r)\Phi(\phi)$ then reduce the above equation to two separate equations for the functions Φ and R

$$\frac{\mathrm{d}^2\Phi}{\mathrm{d}\phi^2} + m^2\Phi = 0 \tag{18}$$

$$\frac{\mathrm{d}^2 R}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}R}{\mathrm{d}r} - \frac{m^2}{r^2}R + \frac{2\mu}{\hbar^2}\left(\frac{k}{r} - E\right)R = 0 \tag{19}$$

where m is identified as the angular momentum of the system. From the system of ordinary differential equations given in Equations (18) and (19), the energy spectrum can be found as

$$E = -\frac{k^2\mu}{2\hbar^2\left(n + m + 1/2\right)^2} \tag{20}$$

It is seen that if the physical system is the Bohr model of two-dimensional hydrogen-like atom then the angular momentum m must take half-integral values. Hence, the topological structure of a configuration space of a physical system can determine the quantum nature of an observable of the system. This result should be expected in quantum mechanics since we know that the quantum behaviour of a particle depends almost entirely on the configuration of an experiment. If, in a particular experiment, the electron of a hydrogen-like atom is constrained to move in a plane, then the orbital angular momentum of the electron must take half-integral values if we use the Schrödinger equation to study the dynamics of the electron and want to retain the same energy spectrum as the Bohr model. As a consequence, it might seem possible to invoke the result to explain the Stern-Gerlach experiment without the necessity of introduction of spin into the quantum theory.

3. Dirac Real Equation

In this section we show that Dirac equation for a free particle can be used to describe the state of a fluid of the quantum particle formulated in the theory of classical fluids. For free Dirac quantum particles, Equation (13) reduces to

$$\left(i\gamma^{\mu}\partial_{\mu}-m\right)\psi=0 \tag{21}$$

By expanding Equation (21) using the matrices γ_i given in Equation (6), we obtain

$$-\frac{\partial\psi_1}{\partial t}-im\psi_1=\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\psi_4+\frac{\partial\psi_3}{\partial z} \tag{22}$$

$$-\frac{\partial\psi_2}{\partial t}-im\psi_2=\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\psi_3-\frac{\partial\psi_4}{\partial z} \tag{23}$$

$$\frac{\partial\psi_3}{\partial t}-im\psi_3=\left(-\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\psi_2-\frac{\partial\psi_1}{\partial z} \tag{24}$$

$$\frac{\partial\psi_4}{\partial t}-im\psi_4=\left(-\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\psi_1+\frac{\partial\psi_2}{\partial z} \tag{25}$$

First, it is observed that with the form of the field equations given in Equations (22)-(25), we may interpret that the change of the field $\left(\psi_1,\psi_2\right)$ with respect to time generates the field $\left(\psi_3,\psi_4\right)$, similar to the case of Maxwell field equations in which the change of the electric field generates the magnetic field. With this observation it may be suggested that, like the Maxwell electromagnetic field which is composed of two essentially different physical fields, the Dirac field of massive particles may also be viewed as being composed of two different physical fields, namely the field $\left(\psi_1,\psi_2\right)$, which plays the role of the electric field in Maxwell field equations, and the field $\left(\psi_3,\psi_4\right)$, which plays the role of the magnetic field. The similarity between Maxwell field equations and Dirac field equations can be carried further by showing that it is possible to reformulate Dirac equation as a system of real equations. When we formulate Maxwell field equations from a system of linear first order partial differential equations we rewrite the original Maxwell field equations from a vector form to a system of first order partial differential equations by equating the corresponding terms of the vectorial equations [9]. Now, since, in principle, a complex quantity is equivalent to a vector quantity therefore in order to form a system of real equations from Dirac complex field equations we equate the real parts with the real parts and the imaginary parts with the imaginary parts. In this case Dirac equation given in Equations (22)-(25) can be rewritten as a system of real equations as follows

$$-\frac{\partial\psi_1}{\partial t}=\frac{\partial\psi_4}{\partial x}+\frac{\partial\psi_3}{\partial z} \tag{26}$$

$$-\frac{\partial\psi_2}{\partial t}=\frac{\partial\psi_3}{\partial x}-\frac{\partial\psi_4}{\partial z} \tag{27}$$

$$-\frac{\partial\psi_3}{\partial t}=\frac{\partial\psi_2}{\partial x}+\frac{\partial\psi_1}{\partial z} \tag{28}$$

$$-\frac{\partial\psi_4}{\partial t}=\frac{\partial\psi_1}{\partial x}-\frac{\partial\psi_2}{\partial z} \tag{29}$$

$$\frac{\partial\psi_4}{\partial y}=m\psi_1 \tag{30}$$

$$\frac{\partial\psi_3}{\partial y}=-m\psi_2 \tag{31}$$

$$\frac{\partial\psi_2}{\partial y}=-m\psi_3 \tag{32}$$

$$\frac{\partial\psi_1}{\partial y}=m\psi_4 \tag{33}$$

If the wavefunction ψ satisfies Dirac field equations given in Equations (26)-(33) then we can derive the following system of equations for all components

$$\frac{\partial^2\psi_i}{\partial^2 y}-m^2\psi_i=0 \tag{34}$$

$$\frac{\partial^2\psi_i}{\partial t^2}-\frac{\partial^2\psi_i}{\partial x^2}-\frac{\partial^2\psi_i}{\partial z^2}=0 \tag{35}$$

Solutions to Equation (34) are

$$\psi_i=c_{1i}\left(x,z\right)\mathrm{e}^{my}+c_{2i}\left(x,z\right)\mathrm{e}^{-my} \tag{36}$$

where c_{1i} and c_{2i} are undetermined functions of $\left(x,z\right)$. The solutions given in Equation (36) give a distribution of a physical quantity along the y-axis. On the other hand, Equation (35) can be used to describe the dynamics, for example, of a vibrating membrane in the $\left(x,z\right)$-plane. If the membrane is a circular membrane of radius a then the domain D is given as $D=\left\{x^2+z^2<a^2\right\}$. In the polar coordinates given in terms of the Cartesian coordinates $\left(x,z\right)$ as $x=r\cos\theta$, $z=r\sin\theta$, the two-dimensional wave equation given in Equation (35) becomes

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2}-\frac{\partial^2\psi}{\partial r^2}-\frac{1}{r}\frac{\partial\psi}{\partial r}-\frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2}=0 \tag{37}$$

The general solution to Equation (37) for the vibrating circular membrane with the condition $\psi=0$ on the boundary of D can be found as [6] [11]

$$\begin{aligned}\psi(r,\theta,t) &= \sum_{m=1}^{\infty} J_0\left(\sqrt{\lambda_{0m}}r\right)\left(C_{0m}\cos\sqrt{\lambda_{0m}}ct + D_{0m}\sin\sqrt{\lambda_{0m}}ct\right) \\ &\quad + \sum_{m,n=1}^{\infty} J_n\left(\sqrt{\lambda_{nm}}r\right)\left(A_{nm}\cos n\theta + B_{nm}\sin n\theta\right) \\ &\quad \times\left(\left(C_{nm}\cos\sqrt{\lambda_{nm}}ct + D_{nm}\sin\sqrt{\lambda_{nm}}ct\right)\right)\end{aligned} \tag{38}$$

where $J_n\left(\sqrt{\lambda_{nm}}r\right)$ is the Bessel function of order n and the quantities A_{nm}, B_{nm}, C_{nm} and D_{nm} can be specified by the initial and boundary conditions. It is also observed that at each moment of time the vibrating membrane appears as a 2D differentiable manifold which is a geometric object whose geometric structure can be constructed using the wavefunction given in Equation (38) as

$$R = \left(2(\psi_{11}\psi_{22} - (\psi_{12})^2\right) \Big/ \left(1 + \psi_1^2 + \psi_2^2\right)^2$$

where $\psi_\mu = \partial\psi/\partial x^\mu$ and $\psi_{\mu\nu} = \partial^2\psi/\partial x^\mu \partial x^\nu$ [25]. Even though elementary particles may have the geometric and topological structures of a 3D differentiable manifold, it is seen from the above descriptions via the Schrödinger wave equation and Dirac equation that they appear as 3D physical objects that embedded in three-dimensional Euclidean space. Interestingly, in the following we will show that the solution given in Equation (38) can be used to describe a standing wave in a fluid due to the motion of two waves in opposite directions. For a steady state in which the system is time-independent, the system of equations given in Equations (26)-(29) reduces to the following system of equations

$$\frac{\partial\psi_2}{\partial x} + \frac{\partial\psi_1}{\partial z} = 0, \frac{\partial\psi_1}{\partial x} - \frac{\partial\psi_2}{\partial z} = 0 \tag{39}$$

$$\frac{\partial\psi_4}{\partial x} + \frac{\partial\psi_3}{\partial z} = 0, \frac{\partial\psi_3}{\partial x} - \frac{\partial\psi_4}{\partial z} = 0 \tag{40}$$

It is observed from Dirac equation for steady states that the field (ψ_1, ψ_2) and the field (ψ_3, ψ_4) satisfy the Cauchy-Riemann equations in the (x, z)-plane. We will now discuss whether it is possible to consider Dirac quantum particles as physical systems which exist in a fluid state as defined in the classical fluid dynamics as substances that retain a definite volume, have the ability to flow and deform continually, hence they can exhibit a wave motion. For references in the next section we will outline the main features in the theory of classical fluids, especially, in two dimensions.

4. Fluid Dynamics in Two-Dimensions

In fluid dynamics, quantities that satisfy the Cauchy-Riemann equations can be identified with the velocity potential and the stream function of an incompressible and irrotational flow [26] [27] [28] [29]. In two-dimensional fluid dynamics, a streamline is a theoretical line that is assumed to be tangential to the instantaneous velocity, therefore there is no flow that can cross the streamline. For a continuous stream of fluid, the streamlines can form continuous lines or closed curves. As an illustration, in the following we will consider a free or potential vortex flow whose streamlines are concentric circles in the (x, z)-plane as shown in Figure 2.

In two-dimensional incompressible flow, the stream function Ψ is defined as a volume flux through a curve given by

$$\Psi = \int \left(v_x \mathrm{d}z - v_z \mathrm{d}x\right) \tag{41}$$

From the definition given in Equation (41), we have $\mathrm{d}\Psi = v_x \mathrm{d}z - v_z \mathrm{d}x$. On the other hand, the total derivative of the stream function Ψ is given by $\mathrm{d}\Psi = (\partial\Psi/\partial x)\mathrm{d}x + (\partial\Psi/\partial z)\mathrm{d}z$, therefore we obtain the following relationships between the velocity components (v_x, v_z) and the stream function Ψ

$$v_x = \frac{\partial\Psi}{\partial z}, v_z = -\frac{\partial\Psi}{\partial x} \tag{42}$$

If the stream function Ψ is defined in terms of polar coordinates (r, θ) as $\Psi = \Psi(r, \theta)$, then we have $\mathrm{d}\Psi = (\partial\Psi/\partial r)\mathrm{d}r + (\partial\Psi/\partial\theta)\mathrm{d}\theta$. As shown in Figure 3, we also have $\mathrm{d}\Psi = v_r (r\mathrm{d}\theta) - v_\theta \mathrm{d}r$.

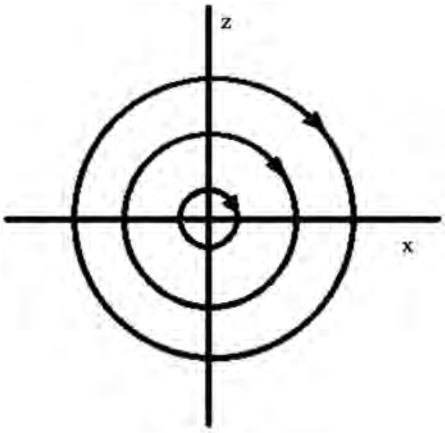


Figure 2. Circular streamlines.

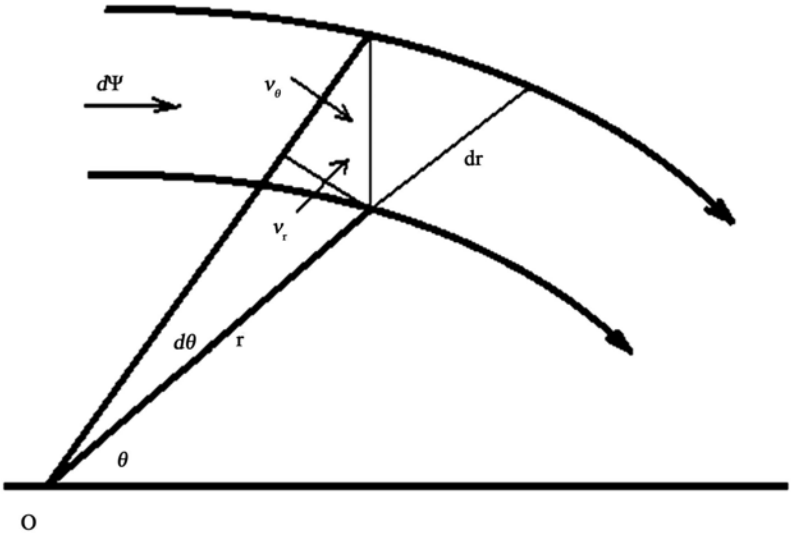


Figure 3. Fluid flow in terms of stream function in polar coordinates.

Therefore we also obtain the following relationships between the velocity components (v_r, v_θ) and the stream function Ψ

$$v_r = \frac{1}{r}\frac{\partial \Psi}{\partial \theta}, v_\theta = -\frac{\partial \Psi}{\partial r} \tag{43}$$

From the definition of the stream function we see that the radial component of the velocity of a vortex flow is equals to zero, $v_r = 0$, since there is no flow that can cross the streamlines. In fluid dynamics, the circulation Γ around a closed curve is a line integral of velocity $\boldsymbol{v}$ defined as

$$\Gamma = \oint \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{s} = \oint v_s \mathrm{d}s \tag{44}$$

where v_s is the tangential velocity. By Stokes' theorem, the circulation Γ is related to the vorticity $\boldsymbol{\omega} = \nabla \times \boldsymbol{v}$ as

$$\Gamma = \oint_{\partial S} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{s} = \iint_S \boldsymbol{\omega} \cdot \mathrm{d}\boldsymbol{S} \tag{45}$$

It is seen from the above equation that the flux of vorticity is the circulation. In particular, for a two-dimensional flow in the (x, z)-plane, the circulation becomes

$$\Gamma = \int_{\mathrm{A}} \left(\frac{\partial v_z}{\partial x} - \frac{\partial v_x}{\partial z} \right) \mathrm{d}x\mathrm{d}z \tag{46}$$

A flow for which the circulation is equal to zero, $\boldsymbol{\omega} = \nabla \times \boldsymbol{v} = 0$, is called a potential or irrotational flow. In two dimensions it is seen from Equation (46) that the condition for potential flow is

$$\frac{\partial v_z}{\partial x} - \frac{\partial v_x}{\partial z} = 0 \tag{47}$$

It should be mentioned that in potential flow we have $\Gamma = \oint_{\partial S} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{s} = 0$ therefore closed streamlines cannot exist in such flow. In general this result prevents us from identifying the components of Dirac equation ψ_1 as closed stream function at the same time identifying ψ_2 as velocity potential. However, it is seen that the result of $\Gamma = \oint \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{s} = 0$ may not be valid if the region of space is multiply-connected since the velocity circulation may not be zero if the closed contour cannot be contracted to a point. This important feature can be discussed further as follows. For a two-dimensional irrotational flow given in polar coordinates, it can be shown that the flow velocity v_θ and the radius r satisfy the following relationship $r v_\theta = C$, where C is a constant. The constant C can be established by using the singularity in the irrotational vortex flow where to velocity becomes infinite at the centre of the vortex with the vorticity is given by the relation $\omega = -(\partial v_\theta / \partial r + v_\theta / r)$. In this case the circulation around a circular streamline can be found as

$$\Gamma = \oint \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{s} = \oint v_s \mathrm{d}s = 2\pi r v_\theta \tag{48}$$

It is also interesting to note that in the Bohr model of the hydrogen atom in which the electron is assumed to move around the nucleus in stationary circular orbits with $v_\theta = v$ then the angular momentum is quantised as $mrv = n\hbar$. If we now also assume that Dirac quantum particles are in fluid states whose circulation Γ is also quantised as the angular momentum then we obtain the following quantisation for the circulation

$$\Gamma = 2\pi r v = \frac{2\pi n\hbar}{m} = \frac{nh}{m} \tag{49}$$

Since $v_r = 0$, the stream function $\Psi = \Psi(r,\theta)$ can be obtained as follows

$$\Psi = -\int v_\theta \mathrm{d}r = -\int \frac{\Gamma}{2\pi r}\mathrm{d}r = -\frac{\Gamma}{2\pi}\ln(k_1 r) = -\frac{n\hbar}{m}\ln(k_1 r) \tag{50}$$

where k_1 is an undetermined constant.

In fluid dynamics, another important concept that is connected with an irrotational flow is the concept of the velocity potential Φ which is defined in the (x,z)-plane as

$$\Phi = \int v_s \mathrm{d}s = \int v_x \mathrm{d}x + v_z \mathrm{d}z \tag{51}$$

It is seen from Equation (51) that the velocity components can be expressed in terms of the velocity potential as follows

$$v_x = \frac{\partial\Phi}{\partial x}, v_z = \frac{\partial\Phi}{\partial z} \tag{52}$$

In polar coordinates (r,θ), the velocity potential and its relationship with the velocity components are given as

$$\Phi = \int v_s \mathrm{d}s = \int v_r \mathrm{d}r + v_\theta r \mathrm{d}\theta, \tag{53}$$

$$v_r = \frac{\partial\Phi}{\partial r}, v_\theta = \frac{1}{r}\frac{\partial\Phi}{\partial\theta} \tag{54}$$

Similarly, the velocity potential Φ can also be obtained using the relation $\mathrm{d}\Phi = r v_\theta \mathrm{d}\theta + v_r \mathrm{d}r$ as

$$\Phi = \int r v_\theta \mathrm{d}\theta = \int r\left(\frac{\Gamma}{2\pi r}\right)\mathrm{d}\theta = \int r\left(\frac{n\hbar}{mr}\right)\mathrm{d}\theta = \frac{n\hbar}{m}\theta + k_2 \tag{55}$$

where k_2 is an undetermined constant. From the relationships given in Equations (42) and (52) we then obtain the Cauchy-Riemann equations

$$\frac{\partial\Phi}{\partial x} + \frac{\partial\Psi}{\partial z} = 0, \frac{\partial\Psi}{\partial x} - \frac{\partial\Phi}{\partial z} = 0 \tag{56}$$

5. Fluid state of Dirac Quantum Particles

By comparing Equation (56) to Dirac equations given in Equations (39) and (40), the field (ψ_1,ψ_2) may be identified as the stream function and the velocity potential of one fluid flow and the field (ψ_3,ψ_4) with another fluid flow. However, the main problem that we want to deal with now is whether the two fields (ψ_1,ψ_2) and (ψ_3,ψ_4) are connected and, most importantly, how such connection would lead to the prospect of using them to describe a Dirac quantum particle as a standing wave. In the following we will show that in fact this is the case by using the relationships between the components of these two fields given in Equations (30)-(33). For convenience we rewrite these equations as follows

$$\frac{\partial \psi_4}{\partial y} = m\psi_1, \frac{\partial \psi_1}{\partial y} = m\psi_4 \tag{57}$$

$$\frac{\partial \psi_3}{\partial y} = -m\psi_2, \frac{\partial \psi_2}{\partial y} = -m\psi_3 \tag{58}$$

$$\frac{\partial^2 \psi_i}{\partial^2 y} - m^2 \psi_i = 0 \quad \text{for } i = 1,2,3,4 \tag{59}$$

If the physical quantity m, which is identified with the inertial mass of a quantum particle, is assumed to be positive, $m > 0$, then it is observed that it is possible to describe the physical structure of a Dirac quantum particle as a spinning top if we consider solutions to Equation (59) as hybrid functions of the form

$$\psi_i = \begin{cases} c_{1i}(x,z)\mathrm{e}^{my} & \text{for } y < 0 \\ c_{2i}(x,z)\mathrm{e}^{-my} & \text{for } y \geq 0 \end{cases} \tag{60}$$

For simplicity, instead of the hybrid form given in Equation (60), in the following we will show only for the case in which $y \geq 0$ since similar results can be obtained for which $y < 0$. The solutions given in Equation (60) can be rewritten in the following forms

$$\psi_1 = c_{21}(x,z)\mathrm{e}^{-my}, \psi_2 = c_{22}(x,z)\mathrm{e}^{-my} \tag{61}$$

$$\psi_3 = c_{23}(x,z)\mathrm{e}^{-my}, \psi_4 = c_{24}(x,z)\mathrm{e}^{-my} \tag{62}$$

Using the equations given in Equations (57) and (58), we further obtain the conditions $c_{24} = -c_{21}$ and $c_{23} = c_{22}$. If we write $c_{21} = f(x,z)$ and $c_{22} = g(x,z)$ then we have

$$\psi_1 = f(x,z)\mathrm{e}^{-my}, \psi_2 = g(x,z)\mathrm{e}^{-my} \tag{63}$$

$$\psi_3 = g(x,z)\mathrm{e}^{-my}, \psi_4 = -f(x,z)\mathrm{e}^{-my} \tag{64}$$

From the above forms of solutions given to the components ψ_i of the wavefunction ψ we can show how a standing wave can be established from the superposition of a wave associated with the field (ψ_1,ψ_2) and a wave associated with the field (ψ_3,ψ_4). Let $\psi_1 = f(x,z)\mathrm{e}^{-my}$ be identified with the velocity potential and $\psi_2 = g(x,z)\mathrm{e}^{-my}$ with the stream function of one fluid flow. Now we have two different descriptions that can be given to the field (ψ_3,ψ_4). If we identify the component $\psi_3 = g(x,z)\mathrm{e}^{-my}$ with the velocity potential and $\psi_4 = -f(x,z)\mathrm{e}^{-my}$ with the stream function of another fluid flow then we have the stream function of the first flow equals the velocity potential of the second flow, and the stream function of the second flow is a reflection of the velocity of the first flow. Even though this kind of identification may be used to describe a particular type of fluid flow of Dirac quantum particles, it does not give rise to the physical structure that we are looking for, that is a standing wave. However, if we now identify the component $\psi_3 = g\left(x,z\right)\mathrm{e}^{-my}$ with the stream function and $\psi_4 = -f\left(x,z\right)\mathrm{e}^{-my}$ with the velocity potential of the second flow then the two flows are identical except for their flow directions, which are opposite to each other, and in fact this is what we want to obtain because they can form a required standing wave. It is also observed that for a steady motion a Dirac quantum particle does not exhibit a wave motion even though it has the potential to establish a wave within its physical structure. Therefore, without an external disturbance a Dirac quantum particle may be considered as a classical particle defined in classical physics. Furthermore, we may also speculate that the two opposite fluid flows associated with the physical structure of a Dirac quantum particle may be related to the concept of spin-half that is introduced into quantum mechanics.

6. Conclusion

In our previous works on spacetime structures of quantum particles, we suggest that all quantum particles are formed from mass points which are joined together by contact forces, which is a consequence of viewing quantum particles as CW-complexes. Being identified with differentiable manifolds, quantum particles therefore should be endowed with geometric and topological structures of differentiable manifolds and their motion should be described as isometric embeddings in higher Euclidean space. In particular, we show that quantum particles may have the geometric and topological structures of a 3D differentiable manifold which can be described as standing waves which are solutions to the Schrödinger wave equation and Dirac equation. In this work we have extended our previous discussions by showing that Dirac equation can be used to describe quantum particles as composite structures that are in a fluid state in which the components of the wavefunction can be identified with the stream function and the velocity potential of a potential flow formulated in the theory of classical fluids. With this fluid composition, physically, Dirac quantum particles can manifest as standing waves which are the result of the superposition of two fluid flows moving in opposite directions. However, for a steady motion, a quantum particle whose physical structure is constructed in terms of Dirac equation does not exhibit a wave motion even though it has the potential to establish a wave within its physical structure. Therefore, if there are no external fields acting on it, a Dirac quantum particle may be considered as a classical particle defined in classical physics. It is also noted from the fact that there are two identical fluid flows in opposite directions within their physical structures, the fluid state model of Dirac quantum particles can be invoked to explain why fermions are spin-half particles as discussed in Section 2.

Acknowledgements

We would like to thank Vanessa Cheng and Jane Gao of the administration of JMP for their editorial advice during the preparation of this work.

Conflicts of Interest

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

References

[1] Ho, V.B. (2018) *International Journal of Physics*, **6**, 105-115.

[2] Ho, V.B. (2018) *GJSFR-A*, **18**, 37-58.

[3] Ho, V.B. (2018) *International Journal of Physics*, **6**, 139-146.

[4] Einstein, A. (1952) The Principle of Relativity. Dover Publications, New York.

[5] Hamilton, R.S. (1982) *Journal of Differential Geometry*, **17**, 255-306. https://doi.org/10.4310/jdg/1214436922

[6] Strauss, W.A. (1992) Partial Differential Equation. John Wiley & Sons, Inc., New York.

[7] Cao, H.D. and Zhu, X.P. (2006) *Asian Journal of Mathematics*, **10**, 165-492. https://doi.org/10.4310/AJM.2006.v10.n2.a2

[8] Ho, V.B. (2017) A Derivation of Dirac Equation from a General System of Linear First Order Partial Differential Equations. Preprint, ResearchGate, viXra 1712.0404v1.

[9] Ho, V.B. (2018) Formulation of Maxwell Field Equations from a General System of Linear First Order Partial Differential Equations. Preprint, ResearchGate, viXra 1802.0055v1.

[10] Melshko, S.V. (2005) Methods for Constructing Exact Solutions of Partial Differential Equations. Springer Science & Business Media, Inc.

[11] Sobolev, S.L. (1964) Partial Differential Equations of Mathematical Physics. Dover Publications, Inc, New York.

[12] Dirac, P.A.M. (1928) *Proceedings of the Royal Society A*: *Mathematical*, *Physical and Engineering Sciences*, **117**, 610-624. https://doi.org/10.1098/rspa.1928.0023

[13] Jackson, J.D. (1975) Classical Electrodynamics. John Wiley & Sons, New York.

[14] Landau, L.D. and Lifshitz, E.M. (1987) The Classical Theory of Fields. Pergamon Press, Sydney.

[15] Ho, V.B. (2018) Topological Transformation of Quantum Dynamics. Preprint, ResearchGate, viXra 1810.0300v1.

[16] Bohr, N. (1913) *Philosophical Magazine*, **26**, 1-25.

[17] De Broglie, L. (1925) *Annales de Physique*, **10**, 22-128.

[18] Lachieze-Rey, M. and Luminet, J.P. (2003) Cosmic Topology.

[19] Ho, V.B. (2018) Formulation of Dirac Equation for an Arbitrary Field from a System of Linear First Order Partial Differential Equations.

[20] Bransden, B.H. and Joachain, C.J. (1986) Physics of Atoms and Molecules. Longman Scientific & Technical, New York.

[21] Ho, V.B. (1994) *Journal of Physics A*: *Mathematical and General*, **27**, 6237-6241. https://doi.org/10.1088/0305-4470/27/18/031

[22] Ho, V.B. and Morgan, M.J. (1996) *Journal of Physics A*: *Mathematical and General*, **29**, 1497. https://doi.org/10.1088/0305-4470/29/7/019

[23] Ho, V.B. (1996) Geometrical and Topological Methods in Classical and Quantum Physics. PhD Thesis, Monash University, Clayton.

[24] Balachandran, A.P., Marmo, G., Skagerstam, B.S. and Stern, A. (1991) Classical Topology and Quantum States. World Scientific, Singapore.

[25] Kreyszig, E. (1975) Introduction to Differential Geometry and Riemannian Geometry. University of Toronto Press, Toronto.

[26] Landau, L.D. and Lifshitz, E.M. (1989) Fluid Mechanics. Pergamon Press, Sydney.

[27] Douglas, J.F., Gasiorek, J.M. and Swaffield, J.A. (1990) Fluid Mechanics. Longman Scientific & Technical, New York.

[28] Batchelor, G.K. (1988) An Introduction to Fluid Dynamics. Cambridge University Press, Melbourne.

[29] White, F.M. (2011) Fluid Mechanics. McGraw Hill, New York.

Scientific Research Publishing

Journal of Modern Physics, 2018, 9, 2420-2436
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Representation of the Basic Laws of Thermodynamics in Quantum Mechanics

Akira Suzuki[1], **Hisao Taira**[2]

[1]Faculty of Science, Tokyo University of Science, Shinjuku-ku, Tokyo, Japan
[2]Faculty of Education, Hokkaido University of Education, Kita-ku, Sapporo, Japan
Email: asuzuki@rs.kagu.tus.ac.jp, taira.hisao@s.hokkyodai.ac.jp

How to cite this paper: Suzuki, A. and Taira, H. (2018) Representation of the Basic Laws of Thermodynamics in Quantum Mechanics. *Journal of Modern Physics*, **9**, 2420-2436.
https://doi.org/10.4236/jmp.2018.914155

Received: September 10, 2018
Accepted: December 11, 2018
Published: December 14, 2018

Abstract

We propose a representation of the basic laws, namely the zeroth, first, second and third law, in quantum thermodynamics. The zeroth law is represented by some parameters (Θ 's) that specify respective quantum states. The parameters are the elements of thermodynamic state space $\mathcal{M}_{\Theta}$. The introduction of such parameters is based on a probabilistic nature of quantum theory. A quantum analog of the first law can be established by utilizing these parameters. The notion of heat in quantum systems is clarified from the probabilistic point of view in quantum theory. The representation of the second law can be naturally described in terms of these parameters introduced for the respective quantum systems. In obtaining the representation of quantum thermodynamics, consistency between quantum theory and classical thermodynamics should have been preserved throughout our formulation of quantum thermodynamics. After establishing the representation of the second law, the third law is discussed briefly. The relationship between thermodynamic temperatures and the parameters in $\mathcal{M}_{\Theta}$ is also discussed.

Keywords

Basic Laws of Thermodynamics, Thermodynamic State Space, Probabilistic Nature of Quantum Theory, Notion of Heat, Entropy Principle, Adiabatic Accessibility

1. Introduction

Thermodynamics is one of the theories which have high universality since thermodynamics as itself has been unchanged even if we now have a well-developed quantum theory. Classical thermodynamics has been well established by different approaches [1] [2] [3]. Lieb and Yngvason made the mathematical structure transparent by axiomatic approach [3]. Thermodynamics is a theory not only for classical but for quantum systems. Above all, the theoretical importance of thermodynamical consideration in quantum systems (quantum thermodynamics) is emphasized in text books [4] [5]. When we consider the thermodynamics for quantum systems, the important is the change in entropy since entropy is a constant of motion under the unitary transformation generated by a system Hamiltonian [6] [7]. A quantum heat engine has been investigated theoretically [8] [9]. Bender *et al.* studied a quantum Carnot cycle by considering a single quantum mechanical particle confined in a quantum well [9]. In their study, they found that the efficiency is equal to that of the Carnot cycle for classical case and proposed that the internal energy U plays the same role as temperature. It should be however mentioned that in quantum system one cannot describe thermodynamic equibria in terms of a parameter like a temperature as in classical system [3].

This paper deals with the following questions that must be answered: Can thermodynamical laws refer to the variation of states of a system represented by the quantum states such as those states (eigenstates) of the Hamiltonian for a single quantum mechanical particle confined in a quantum well? If thermodynamical laws exist in quantum thermodynamic systems, how can they be expressed? To answer these questions, we need a representation which connects thermodynamic states and quantum states.

In classical thermodynamics, states are represented by points on a state space. A typical example of the state space is just a collection of P (pressure) and V (volume), *i.e.*, a P-V plane. In quantum mechanics, what space should be used in order to describe thermodynamic states for quantum systems? In quantum mechanics, quantum (pure) states are expressed by the elements of a complex Hilbert space $\mathcal{H}$. However, the Hilbert space itself does *not* play the same role as the state space in classical thermodynamics since comparing one state vector with the others in $\mathcal{H}$ must be done by comparing the components of each vector. Thus we start with introducing a set $\mathcal{M}_\Psi$ of the state vectors in $\mathcal{H}$ in order to obtain a suitable set which plays the same role as the state space in classical thermodynamics. That is, we introduce a set $\mathcal{M}_\Theta$ which plays the same role as the state space and derive a correspondence between $\mathcal{M}_\Psi$ and $\mathcal{M}_\Theta$. After that, we will show that the first law of quantum thermodynamics can be described by the elements in $\mathcal{M}_\Theta$ and the internal energy of a quantum system can be described as a function on $\mathcal{M}_\Theta$. We will also discuss the relation between those parameters Θ's in $\mathcal{M}_\Theta$ and thermodynamic temperatures.

The first law of thermodynamics is a law of conservation of energy and states the equivalence of heat and work. We will discuss the equivalence between work and heat in quantum thermodynamics. We assume that the energy of a system (*i.e.*, the internal energy) is given by the expectation value of the Hamiltonian $\mathscr{H}$: $U = \langle \mathscr{H} \rangle = \sum_i p_i E_i$. In this expression, E_i is the outcome of the expected energy state corresponding to a definite probability p_i in a specified maximal test. Indeed, this probabilistic nature of quantum system plays a key role to establish a representation of the first law. Differentiating U formally, we obtain the expression, $dU = \sum_i \left(E_i dp_i + p_i dE_i \right)$. The first term $\sum_i E_i dp_i$ implies there exists a non-mechanical source that induces a change in the internal energy of the system since a change in quantum states is in general determined by the unitary operator which does not change the definite probability (*i.e.*, $dp_i = 0$). The second term $\sum_i p_i dE_i$ implies a mechanical source that induces a change in the internal energy since we can trace the origin of dE_i to an external parameter. As will be shown in Subsec. 3.2, the following identifications, $d'Q = \sum_i E_i dp_i$ and $d'W = \sum_i p_i dE_i$, are justified and are ensured by the existence of respective parameters Θ and L. We will show that the internal energy of quantum system is generally expressed in terms of parameters Θ and L, respectively, describing the equivalence relation among quantum states *and* external parameters. Therefore, the first law of quantum thermodynamics can be uniquely represented by these parameters. Once establishing the representation of the first law, it is worth to investigate the remaining thermodynamical laws (the second and third laws) for a quantum system described by quantum states.

In this paper, we propose a representation of the thermodynamical laws for quantum system in terms of the respective parameters and develop a theory of quantum thermodynamics based on the axiomatic theory of classical thermodynamics by Lieb and Yngvason [3]. In their formulation, the second law refers to the possible adiabatic transition of any two states in a state space.

This paper is organized as follows. In the next section, we present a brief review of classical thermodynamics. In Sec. 3, we state the basic notion of our formulation of quantum thermodynamics, and introduce a thermodynamic state space $\mathcal{M}_\Theta$ and a quantum state space $\mathcal{M}_\Psi$ and discuss the connection between them. In Subsec. 3.1 we show the existence of the zeroth law of quantum thermodynamics in the state space $\mathcal{M}_\Theta$. In Subsec. 3.2, the first law of thermodynamics and an adiabatic process are discussed. In Subsec. 3.3, we define entropy and give a representation of the second law, and discuss a relation among the adiabatic transitions, entropy and the term $d'Q$. We refer to the third law in Subsec. 3.4. Finally, we give the results and discussion in Sec. 4.

2. Classical Thermodynamics

There are few approaches in thermodynamics [1] [2] [3]. Lieb and Yngvason's approach is helpful to understand the logical structure of thermodynamics. If thermodynamical laws exist in quantum systems as well as in classical systems, there must be the same logical structure in both systems. According to their formulation, a structure of adiabatic accessibility on a state space (*thermodynamic state space*) is characterized by an entropy inequality, *i.e.*, the second law of thermodynamics. In this section we present a brief review of classical thermodynamics due to Lieb and Yngvason [3]. Thermodynamics is a theory which discusses a transition between equilibrium states. The second law refers to the feasible transitions in adiabatic process.

We start with introducing a formulation of the axiomatic thermodynamics proposed by Lieb and Yngvason [3]. In their formulation, the second law of thermodynamics is represented by the entropy principle.

Entropy principle: There is a real-valued function on all states of all systems (including compound systems), called *entropy* and denoted by S. Entropy has the following properties:

- *Monotonicity*: When X and Y are comparable states[1], then
$$X \prec Y \text{ if and only if } S(X) \leq S(Y). \tag{1}$$
- *Additivity*: If X and Y are states of some (possibly different) systems and if (X,Y) denotes the corresponding state in the composition of the two systems, then the entropy is additive for these states, *i.e.*,
$$S(X,Y) = S(X) + S(Y). \tag{2}$$
- *Extensivity*: S is extensive, *i.e.*, for each $t > 0$ and for each state X and its scaled copy tX,
$$S(tX) = tS(X). \tag{3}$$

It should be noted that entropy is determined by the physical (or thermodynamic) state of the system. In the entropy principle, X and Y (*e.g.*, energy and volume) describe equilibrium states and are the elements of a state space (denoted by Γ). A system is then represented by the state space Γ on which a relation "$\prec$" of adiabatic accessibility is defined. The definition of *adiabatic accessibility* is as follows:

Adiabatic accessibility: A process whose only effect on the surroundings is exchange of energy with a mechanical source. This means that as a state arrives at new one, a state of surroundings is the same as before, in other words, the device returns to its initial state at the end of the process.

Lieb and Yngvason [3] proved that *existence* and *uniqueness* of entropy are equivalent to certain simple properties of a relation "$\prec$" (**A1~A6**) and a comparison hypothesis (**Ch**):

A1. *Reflexivity*: $X \overset{\text{A}}{\sim} X$.

A2. *Transitivity*: $X \prec Y$ and $Y \prec Z$ implies $X \prec Z$.

A3. *Consistency*: $X \prec X'$ and $Y \prec Y'$ implies $(X,Y) \prec (X',Y')$.

A4. *Scaling invariance*: If $X \prec Y$, then $tX \prec tY$ for $\forall t > 0$.

A5. *Splitting* and *Recombination*: For $0 < t < 1$, $X \overset{\text{A}}{\sim} (tX, (1-t)X)$.

A6. *Stability*: If, for some pair of states, X and Y, $(X, \epsilon Z_0) \prec (Y, \epsilon Z_1)$ holds for a sequence of ϵ's tendency to zero and some states Z_0 and Z_1, then $X \prec Y$.

---

[1] The word "comparable" used in this paper means that any two states, X and Y, in the same state space hold the relation, either $X \prec Y$ or $Y \prec X$, with respect to the relation "$\prec$". In this context, X and Y are comparable and these states are called *comparable states*.

Ch. *Comparison hypothesis* : The **Ch** holds for a state space Γ if any two states X and Y on the space are comparable states, *i.e.*, $X \prec Y$ or $Y \prec X$.

In the axiom **A1**, the symbol $\overset{\mathrm{A}}{\sim}$ denotes that two states X and Y are *adiabatic* equivalent; It describes a situation where both of the relations, $X \prec Y$ and $Y \prec X$, hold. It should be noted that the axioms, **A3**, **A5** and **A6**, are defined on the product of state space $\Gamma \times \Gamma$, where $(X,Y) \in \Gamma \times \Gamma$. The **Ch** asserts that any two states on the *same* state space are comparable. Generally, the structure on the state space Γ is determined by the axioms (**A1~A6**) and the comparison hypothesis (**Ch**) under the condition of adiabatic accessibility.

Let us consider a meaning of the entropy principle. Let $X, X', Y, Y', \cdots$ be the elements of the state space Γ. Imagine that we have a list of all possible pairs of states X , Y such that Y is adiabatic accessible from X . The foundation of thermodynamics and the essence of the second law are that this list, X and Y , such as $X \prec Y$, can be simply encoded by the entropy function S defined on a set of all states of systems (including compound systems). This means that Y is adiabatic accessible from X , *i.e.*, $X \prec Y$ if and only if $S(X) \le S(Y)$ (*entropy inequality*). The entropy function should be kept consistency with the structure of the state space Γ characterized by **A1~A6** and **Ch**. Thus, we can characterize the structure based on the definition of adiabatic accessibility on the state space Γ by using the entropy inequality. Combining the axioms (**A1~A6**) and the Equation (2), one can describe the entropy principle for systems including a compound system.

Let us consider a compound system in which X , X' and Y , Y' are the states of system A and system B, respectively. In this case, the *entropy principle* is mathematically expressed as follows:

$$(X,Y) \prec (X',Y') \text{ if and only if } S(X)+S(Y) \le S(X')+S(Y'). \tag{4}$$

Note that all states (X',Y') such that $X \prec X'$ and $Y \prec Y'$ are adiabatically accessible from (X,Y). It is then important to notice that (X',Y') can be adiabatically accessible from (X,Y) even if X' is *not* adiabatically accessible from X . In such a case, entropy increase, $S(Y')-S(Y)$, in the process compensates for a loss, $S(X')-S(X)$, so as to satisfy the statement (4). Therefore, the inequality, $S(X)+S(Y) \le S(X')+S(Y')$, characterizes the possible adiabatic transitions for the compound system even when $S(X) \ge S(X')$. It means that it is sufficient to know the entropy of each part of the compound system in order to decide which transition is feasible due to the interactions between the two subsystems.

For later use we write the entropy principle [the statement (4)] in terms of U and V , where U , V denote the internal energy and the volume of a system, respectively. Putting $X=(U_{\mathrm{A}},V_{\mathrm{A}})$ and $Y=(U_{\mathrm{B}},V_{\mathrm{B}})$, one obtains from the statement (4):

$$\begin{gathered}\left(\left(U_{\mathrm{A}},V_{\mathrm{A}}\right),\left(U_{\mathrm{B}},V_{\mathrm{B}}\right)\right)\prec\left(\left(U'_{\mathrm{A}},V'_{\mathrm{A}}\right),\left(U'_{\mathrm{B}},V'_{\mathrm{B}}\right)\right)\\ \text{if and only if } S\left(U_{\mathrm{A}},V_{\mathrm{A}}\right)+S\left(U_{\mathrm{B}},V_{\mathrm{B}}\right)\le S\left(U'_{\mathrm{A}},V'_{\mathrm{A}}\right)+S\left(U'_{\mathrm{B}},V'_{\mathrm{B}}\right).\end{gathered}\tag{5}$$

It should be noted that the state of the compound system composed of system A and system B is described by $\left(U_{\mathrm{A}},U_{\mathrm{B}}\right)$ only in the case where the volume is invariant during the process. The statement (5) makes sense in the case where X is an extensive variable. However, there exists a particular case in which X is an intensive variable.

One of the aims in this paper is to obtain a representation of entropy inequality for quantum system corresponding to the statement (4).

3. Quantum Thermodynamics

In order to obtain the representation of the zeroth, first, second, and third laws for quantum thermodynamics, we have to introduce a *state space* in order to describe thermodynamic states of quantum system. In the previous section, we have seen that the thermodynamic states of classical system denoted by capital Roman letters, X,Y,Z, etc. defined as the elements of state space Γ satisfy certain simple properties of the relation "$\prec$" (**A1~A6**) and the comparison hypothesis (**Ch**). From the mathematical point of view, we expect that the thermodynamic structure of quantum thermodynamics should also have the same structure as that of classical thermodynamics.

In order to develop a representation of quantum thermodynamics, we must introduce a *thermodynamic state space* for quantum system since in quantum theory, quantum system is described by the complex Hilbert space $\mathcal{H}$ and the states of quantum system are in general described by the elements in $\mathcal{H}$: $\left|\Psi_{\alpha}\right\rangle$, $\left|\Psi_{\beta}\right\rangle$, etc. Here, Greek letter (subscript) denotes the label of respective states of quantum system. In the following, we use a symbol $\mathcal{M}_{\Psi}$ for *quantum state space* and a symbol $\mathcal{M}_{\Theta}$ (instead of Γ in classical case) to represent *thermodynamic state space* for quantum system. Our aim in this section is to show the relation between $\mathcal{M}_{\Psi}$ and $\mathcal{M}_{\Theta}$ in order to obtain the representation of quantum thermodynamics. To define the state space for quantum thermodynamics, namely *thermodynamic state space* in quantum systems, we have to establish the relation between $\mathcal{M}_{\Psi}$ *and* $\mathcal{M}_{\Theta}$. In this section we shall introduce a thermodynamic state space $\mathcal{M}_{\Theta}$ and obtain the representation of the zeroth law of quantum thermodynamics.

We first define a set $\mathcal{M}_{\Psi}$ as a *quantum state space*:

$$\mathcal{M}_{\Psi}:=\left\{\left|\Psi_{\alpha}\right\rangle,\ \left|\Psi_{\beta}\right\rangle,\ \cdots\right\}.\tag{6}$$

The set $\mathcal{M}_{\Psi}$ includes those elements which are linear combinations of the elements in $\mathcal{M}_{\Psi}$; *e.g.*, $\left|\Psi_{\gamma}\right\rangle=\left|\Psi_{\epsilon}\right\rangle+c\left|\Psi_{\zeta}\right\rangle$, where c is a complex number (a relative phase between $\left|\Psi_{\epsilon}\right\rangle$ and $\left|\Psi_{\zeta}\right\rangle$). The importance of this statement is well recognized in quantum theory [5]. All state vectors are thus found in the set $\mathcal{M}_{\Psi}$ and they represent respective quantum states of the system. Now we would like to relate each element of thermodynamic states in $\mathcal{M}_\Theta$ to those state vectors in $\mathcal{M}_\Psi$.

To find out the representation of the zeroth law associated with the property of **A1** for quantum thermodynamics, we introduce a set $\mathcal{M}_\Theta$ as a *thermodynamic state space*. It should be noted that the elements in $\mathcal{M}_\Theta := \{\Theta_\alpha, \Theta_\beta, \Theta_\gamma, \cdots\}$ have to be *comparable* each other in a context of the zeroth law of classical thermodynamics. As will be shown below, introducing these parameters Θ's enables us to compare the quantum states in $\mathcal{M}_\Psi$ in thermodynamic sense.

We start with discussing the existence and uniqueness of Θ for quantum state $|\Psi\rangle$ representing a quantum system. Let us introduce the *propositions of quantum theory*, **Q1** and **Q2**: [5].

Q1. A state is characterized by the probabilities of the various outcomes of every conceivable test.

Q2. If a quantum system is prepared in such a way that it certainly yields a predictable outcome in a specified maximal test[2], the various outcomes of any other tests also have definite probabilities. In particular, these probabilities do not depend on the details of the procedure employed for preparing the quantum system. Therefore, the quantum system so prepared yields a specific outcome in the given maximal test. It should be noted that the quantum system prepared in such a way is said to be in a *pure* state.

Any complete orthogonal basis, $|i\rangle$, $i = 1, 2, \cdots, i, \cdots$, represents a realizable maximal test. Therefore one can obtain the *definite* probability p_i for state *i* from a probability amplitude: [10]

$$a_i = \langle i | \Psi \rangle. \tag{7}$$

It should be noted that the definite probability is given by $p_i = |a_i|^2$.

Now we introduce a *lemma* (**L1**) on the existence of parameter Θ_j $(\mathrm{j} = \alpha, \beta, \gamma, \cdots)$:

L1. There exist parameters $\Theta_\alpha, \Theta_\beta, \cdots$ in $\mathcal{M}_\Theta$ for respective state vectors, $|\Psi_\alpha\rangle, |\Psi_\beta\rangle, \cdots$ in $\mathcal{M}_\Psi$.

Proof of L1. We treat a case for a label α. Other cases, $\beta, \gamma, \cdots$, could be proved in the same way. From Equation (7), one can obtain a sequence $\left\{|a_1^\alpha|^2, |a_2^\alpha|^2, \cdots, |a_i^\alpha|^2, \cdots\right\}$. By normalizing a state vector $|\Psi_\alpha\rangle$, the sequence satisfies the condition $\sum_{i=1} |a_i^\alpha|^2 = 1$ and it is clear that $0 \le |a_i^\alpha|^2 \le 1$. Then each element of the sequence describes a definite probability for the respective state *i*. Hereafter, we shall omit superscript α for simplicity. Let us introduce a probability function $P(\Theta; x)$ so as to satisfy $P(\Theta; x = i) = |a_i|^2$ for any *i*. In

---

[2] Maximal test is defined as follows: Let *N* (assumed to be finite for simplicity) be the maximum number of different outcomes obtainable in a test of a given quantum system. Then, any tests that have exactly *N* different outcomes are called *maximal* (complete). Such tests are called *maximal test*.

$P(\Theta;x)$, Θ is a parameter and x is a random variable. Therefore, the parameter Θ_j $(\mathrm{j}=\alpha,\beta,\gamma,\cdots)$ exists for the representation of a state vector $|\Psi_\mathrm{j}\rangle$. □

The **L1** ensures the existence of parameters Θ's which correspond to respective state vectors $|\Psi_\mathrm{j}\rangle$'s in $\mathcal{M}_\Psi$ through probability functions $P(\Theta_\mathrm{j};x)$'s. It is noted that a parameter is thereby found in each probability function for respective labels, $\mathrm{j}=\alpha,\beta,\gamma,\cdots$. Since Greek letters refer to the respective quantum states, $|\Psi_\mathrm{j}\rangle$ corresponds to Θ_j $(\mathrm{j}=\alpha,\beta,\gamma,\cdots)$. In **L1**, we saw that there exists *unique* parameters Θ's assigned for *every* state vector in $\mathcal{M}_\Psi$. However, in order to represent the zeroth law in terms of Θ's, the elements in $\mathcal{M}_\Theta$ must be mathematically comparable each other. Thus, it is required to utilize the *same* maximal test in order to obtain the equivalence relation for the element Θ_j from the corresponding equivalence relation for the state vector $|\Psi_\mathrm{j}\rangle$ for $\mathrm{j}=\alpha,\beta,\cdots$.

Adopting this requirement, it is ensured that those elements in $\mathcal{M}_\Theta$ are comparable each other and the **Ch** holds for Θ_j's once entropy is defined as a function of Θ. Therefore, the zeroth law of quantum thermodynamics can be uniquely established for quantum systems by making use of parameters Θ_j's. In order to prove the existence of the zeroth law and to discuss the parameter Θ, we have to introduce a proposition of quantum theory, **Q3**:

Q3. $|\Psi\rangle$ describes the same state as $c|\Psi\rangle$, where c is a complex number.

Here, we discuss the parameter Θ associated with the proposition **Q3**. Let us introduce a map $f:\mathcal{M}_\Psi \mapsto \mathcal{M}_\Psi^c$, where $\mathcal{M}_\Psi^c$ is a set defined by $\mathcal{M}_\Psi^c := \{c|\Psi_\alpha\rangle, c|\Psi_\beta\rangle,\cdots\}$. From the **Q3**, $\mathcal{M}_\Psi \equiv \mathcal{M}_\Psi^c$ since it is required that the set $\mathcal{M}_\Psi$ is invariant under the scalar transformation f. Therefore, the set $\mathcal{M}_\Theta$ is also invariant under the scalar transformation, *i.e.*, $f:\mathcal{M}_\Theta \mapsto \mathcal{M}_\Theta^c$. This can be described in terms of their elements, *i.e.*, $f:\Theta_\mathrm{j} \mapsto c\Theta_\mathrm{j}$ $(\mathrm{j}=\alpha,\beta,\gamma,\cdots)$. Thermodynamically, this means that the parameter Θ is an *intensive* variable. It should be noticed that *the zeroth law is invariant under the scalar transformation of state vectors*.

Let us derive the thermodynamic state space $\mathcal{M}_\Theta$ from $\mathcal{M}_\Psi$, where the elements in $\mathcal{M}_\Theta$ correspond to respective state vectors in $\mathcal{M}_\Psi$. It should be noted that in order to obtain the corresponding equivalence relation for these elements Θ_j from the equivalence relation for $|\Psi_\mathrm{j}\rangle$, $(\mathrm{j}=\alpha,\beta,\gamma,\cdots)$, we have to use the *same* maximal test. In other words, when we compare the quantum states, those vectors in $\mathcal{M}_\Psi$ representing the states must be *comparable* and hence their components must be described by the *same* complete orthogonal basis. Adopting this requirement, it is ensured that the elements in $\mathcal{M}_\Theta$ are comparable each other and can specify the respective states of quantum system. Though $\mathcal{M}_\Theta$ should be written as $\mathcal{M}_{\Theta,L}$, where L

denotes a label of maximal test, L will be omitted when we consider a fixed L for the sake of simplicity.

Now one can compare two quantum states in thermodynamic sense since those parameters Θ's can be used to describe two or more systems being equivalent. This leads to the zeroth law of quantum thermodynamics.

3.1. The Zeroth Law of Quantum Thermodynamics

Let us prepare three systems, $\mathcal{M}_\Theta^{\mathrm{A}}$, $\mathcal{M}_\Theta^{\mathrm{B}}$ and $\mathcal{M}_\Theta^{\mathrm{C}}$, where superscripts indicate the labels of respective systems. The *equivalence relation* among quantum states is described by

$$\text{If } \Theta_\alpha^{\mathrm{A}} = \Theta_\beta^{\mathrm{B}}, \text{ then } \left|\Psi_\alpha\right\rangle^{\mathrm{A}} \doteq \left|\Psi_\beta\right\rangle^{\mathrm{B}}, \tag{8}$$

where $\Theta_\alpha^{\mathrm{A}} \in \mathcal{M}_\Theta^{\mathrm{A}}$ and $\Theta_\beta^{\mathrm{B}} \in \mathcal{M}_\Theta^{\mathrm{B}}$, respectively. In this relation, a symbol $\doteq$ means that the state in the left-hand side is *equivalent* to the state in the right-hand side. Later this relation (8) reaffirm the zeroth law of thermodynamics.

Proof of the zeroth law. Let $\Theta_\alpha^{\mathrm{A}}$ and $\Theta_\beta^{\mathrm{B}}$ be the elements of $\mathcal{M}_\Theta^{\mathrm{A}}$ and $\mathcal{M}_\Theta^{\mathrm{B}}$, respectively. By **L1**, it is clear that $\Theta_\alpha^{\mathrm{A}} = \Theta_\beta^{\mathrm{B}}$ implies $\left|\Psi_\alpha\right\rangle^{\mathrm{A}} \doteq \left|\Psi_\beta\right\rangle^{\mathrm{B}}$. ☐

We are now in a position to discuss some consequences obtained by introducing the parameter Θ_{j} to specify the corresponding thermodynamic states of quantum systems. It is clear from Equation (8) that the *transitiveness law* holds:

$$\text{If } \Theta_\alpha^{\mathrm{A}} = \Theta_\beta^{\mathrm{B}} \text{ and } \Theta_\beta^{\mathrm{B}} = \Theta_\gamma^{\mathrm{C}}, \text{ then } \left|\Psi_\alpha\right\rangle^{\mathrm{A}} \doteq \left|\Psi_\gamma\right\rangle^{\mathrm{C}}. \tag{9}$$

We have established a representation of the zeroth law of quantum thermodynamics (equivalence relation among quantum states). The zeroth law can be expressed in terms of parameters in $\mathcal{M}_\Theta$.

3.2. The First Law of Quantum Thermodynamics

The first law of thermodynamics is the law of conservation of energy and it assures equivalence between heat and work. It is written as

$$dU = d'Q + d'W, \tag{10}$$

where dU is a small change in the internal energy of a system, $d'Q$ is the heat transferred to the system and $d'W$ is the work done on the system. Though Equation (10) is for classical system, we will show the same expression holds for quantum system.

We consider a quantum system described by Hamiltonian $\mathscr{H}$. Here we only assume that the internal energy of the system, U, is given by the pure state expectation value of Hamiltonian $\mathscr{H}$:[3]

$$U = \langle \mathscr{H} \rangle = \sum_i p_i E_i, \tag{11}$$

---

[3] The energy of an isolated system may be described by the sum of kinetic and potential energies, represented by Hamiltonian.

where p_i denotes a definite probability having the outcome E_i in a specified maximal test for the quantum system. Since the internal energy (11) can be expressed as $U = U(\Theta, L)$, we can obtain the representation of the first law of quantum thermodynamics in terms of Θ's and *L*'s in $\mathcal{M}_{\Theta,L}$ (see below).

Let us consider a small change in the internal energy. Differentiating Equation (11) formally, we obtain

$$dU = \sum_i \left(E_i dp_i + p_i dE_i \right). \tag{12}$$

This indicates that the change in U is originated from the two independent source, *i.e.*, an outcome E_i *and* a definite probability p_i. The two terms in the right-hand side of Equation (12) can be identified with the respective terms of Equation (10) for the expression of the first law of thermodynamics. The first term $\sum_i E_i dp_i$ is a consequence of the change in probabilities. We shall see that it corresponds to the *heat flow* into the system. This implies that the effect of heat transfer is to change the definite probabilities of the various outcome of E_i. The second term $\sum_i p_i dE_i$ relates to the change in the outcome dE_i. This term shows that change in the internal energy partially comes from the change in outcome. We shall show below that this corresponds to the work done on the system, *i.e.*, $d'W$. It must be emphasized that the validity of Equation (12) is ensured by Equations (13) and (19), *i.e.*, existence of parameters, Θ and L, as will be shown shortly.

In order to understand that the second term corresponds to $d'W$, we first note that the work done on the system is related to the variation of an external parameter (*work coordinate*) and the outcome should depend on the size (volume) of the system: $E_i = E_i(V)$. The work done on the system is due to the change of the volume of the system. Accordingly, without loss of generality we ascribe the change of the volume to the change of the work coordinate denoted by L: the energy of state i, *i.e.*, outcome E_i, which is the eigenvalue of the system's Hamiltonian $\mathscr{H}$, depends thus on the work coordinate L: $E_i = E_i(L)$. This expression states that one can identify the work coordinate L with the label of a maximal test introduced in the last section. Since E_i is a function of L, we can write that the change in the outcome E_i is expressed in terms of the work coordinate:

$$dE_i = \frac{\partial E_i(L)}{\partial L} dL. \tag{13}$$

Then

$$\sum_i p_i dE_i = \sum_i p_i \frac{\partial E_i}{\partial L} dL. \tag{14}$$

In Sec. 3, it was shown that $p_i = P(\Theta; x = i)$. This indicates that p_i is a function of Θ. Thus we can write Equation (14) as

$$\sum_i p_i dE_i = \sum_i \frac{\partial}{\partial L}\left(p_i E_i \right) dL = \frac{\partial}{\partial L} U dL. \tag{15}$$

A change in the internal energy of the system is generally related to a “force” defined by

$$F \equiv -\frac{\partial U(L)}{\partial L}, \tag{16}$$

so that Equation (15) and hence Equation (14) can be expressed as

$$\sum_i p_i dE_i = -FdL. \tag{17}$$

This indicates that the term $\sum_i p_i dE_i$ corresponds to the work $d'W$ done on the system and it is generally expressed by[4]

$$d'W = \sum_i p_i dE_i(L) = -FdL, \tag{18}$$

where the definite probability p_i can then be replaced by the probability function $P(\Theta; x = i)$. It is important to notice that Equation (18) expresses the work $d'W$ with respect to the change in L (work coordinate) in a quasi-static process. Hence, the invariance of an external parameter L implies $d'W = 0$. Thereby we have $dU = d'Q$. Then the term $\sum_i E_i dp_i$ corresponds to the change in the internal energy of the system that occurs when *no* work is done; this is what we understand as heat flow.

It should be emphasized that the heat entering the system, $d'Q$, is expressed in terms of the variation of $p_i = P(\Theta; x = i)$ while the work done on the system, $d'W$, is expressed in terms of the variation of $E_i(L)$. Since $p_i = P(\Theta; x = i)$, we can write dp_i as

$$dp_i = \frac{\partial P(\Theta; x = i)}{\partial \Theta} d\Theta, \tag{19}$$

so that $d'Q$ is expressed in terms of Θ:

$$d'Q = \sum_i E_i dp_i = \sum_i E_i \frac{\partial P(\Theta; x = i)}{\partial \Theta} d\Theta. \tag{20}$$

In general, unitary operator generated by the Hamiltonian of a system does *not* change the definite probability p_i. Therefore, the term $E_i dp_i$ representing the heat cannot be explained by the Hamiltonian itself. In the present theory, the effect of heat is expressed as a change in the definite probabilities having various outcomes E_i in the maximal test. It should be noted that an adiabatic process is characterized by the case where $dp_i = 0$ in the definition of heat, see Equation (20). After all, the representation for the heat in the present theory is consistent with a well-known adiabatic equivalence and the first law of thermodynamics for quantum systems can be represented by the elements in $\mathcal{M}_{\Theta,L}$ as seen in Equations (18) and (20).

As a consequence of the previous section (see the proof of **L1**) and the discussion above, the following *corollaries*, **C1** and **C2**, can be drawn:

C1. The internal energy U is specified by the parameter Θ and L:

---

[4] We note that Equation (18) is a quantum analog of the expression $d'W = -PdV$ in classical case.

$$U = U(\Theta, L). \tag{21}$$

Proof of C1. Once a maximal test is chosen for a fixed L, respective outcome E_i $(i = 1, 2, \cdots)$ for the system is uniquely determined and the definite probability p_i is then described as $P(\Theta; x = i)$ by the proof of **L1**. Therefore, it is clear that the internal energy is specified by the parameters Θ and L. ☐

We note that **C1** states the internal energy U can be specified by the parameters, Θ and L. We will omit L in $U(\Theta, L)$ for simplicity when we consider a fixed L.

C2. $d'Q = 0$ implies the consequence of adiabatic equivalence.

Proof of C2. By the proof of **L1**, *no* change in the probability function implies that the absolute values of the expansion coefficients $|a_i|$'s remain the same. This implies p_i remains constant. Thus $d'Q$ is equal to zero throughout the operation (during the process). ☐

In the statement of **C2**, the consequence of adiabatic equivalence is as follows: if the system is isolated, the absolute values of the expansion coefficients $|a_i|$'s would remain constant [11]. It should be noticed that the notion of heat arises *only* when the state (internal energy) of a system changes, where $dU \neq d'W$. As in classical thermodynamics, heat in quantum systems is also defined as a form of energy movement. Once the internal energy of a quantum system is well defined, heat is also well defined. The notion of heat in quantum systems will be discussed further in a separate paper [12].

3.3. The Second Law and Entropy

In this section we will give a definition of entropy to describe the entropy principle (namely, the *second law of thermodynamics*) for quantum systems. The entropy principle states that the adiabatic accessibility of any two states is described by an entropy inequality. Here we should refer to the adiabatic process since the second law treated here is defined for the process. The process is characterized by $d'Q = 0$. This is ensured when $P(\Theta; x)$ remains unchanged throughout the process (see **C1** and argument below). In other words, adiabatic process is a process such that $P(\Theta; x)$ remains unchanged. It should be noted that adiabatic process allows to change a value of L since it only affects work $d'W$. This is consistent with adiabatic process defined by Lieb and Yngvason [3].

Let us define an entropy function as a map from the set $\mathcal{M}_{\Theta,L}$ to a real number R:

$$S : \mathcal{M}_{\Theta,L} \mapsto R. \tag{22}$$

We note that this general definition for entropy can describe all types of entropy functions including well known Boltzmann, Gibbs, and Shannon entropies. The entropy S defined by the map (22) is clearly a state quantity and ensures that S can be defined for *all* states in terms of Θ.

In order to obtain a representation of the second law in terms of Θ for quantum systems, however, it must be shown that determining a parameter Θ as a state variable means to determine an internal energy U as a state quantity. We have already shown that U is specified by Θ's and L's (*i.e.*, the elements of $\mathcal{M}_{\Theta,L}$) as described in **C1** and can be expressed by a function on $\mathcal{M}_{\Theta,L}$. Thus the following *lemma* (**L2**) is established:

L2. There exists U specified by each element of $\mathcal{M}_{\Theta,L}$.

Proof of L2. Without loss of generality, one can consider a fixed maximal test, where outcome of the maximal test is *uniquely* determined: The internal energy is represented by $U=\sum_i p_i E_i=\sum_i P(\Theta_\alpha;x=i)E_i$, whereby Θ_α specifies the internal energy U. Therefore, U can be labeled as U_α. As in the same way, respective internal energies, $U_\beta,U_\gamma,\cdots$, can be specifically specified by $\Theta_\beta,\Theta_\gamma,\cdots$. □

Since the existence of correspondence between an internal energy U and a parameter Θ was established by **L2**, we can obtain one-to-one correspondence between $S(X)$ and $S(\Theta)$. This keeps consistency between an entropy function defined in the entropy principle and the statement (22). Put U instead of X in the statement (1) in Sec. 2 and we finally obtain a representation of the second law of quantum thermodynamics in terms of Θ:

$$\Theta_\alpha \prec \Theta_{\alpha'} \text{ if and only if } S(\Theta_\alpha)\le S(\Theta_{\alpha'}). \tag{23}$$

This describes the entropy principle for quantum system. It should be noted that each Θ can depend on different values of L. Thus, Eq. (23) is able to give the information as to a complete structure in $\mathcal{M}_{\Theta,L}$, which enables us to compare Θ's that depend on different values of L's in a context of adiabatic accessibility (recall that one can compare state vectors when L's are fixed). We should note that two quantum states represented by Θ_α and Θ_β are *adiabatic equivalent* if and only if $S(\Theta_\alpha)=S(\Theta_\beta)$ for a *fixed* work coordinate L (*cf.*, **A1** for classical case and the argument therein).

Now we consider the case corresponding to the statement (4) in Sec. 2. This states that the second law can refer to a *possible adiabatic transition* of a compound system consisting of system A and B when the compound system is thermally isolated from any other environment. From **L2**, we can immediately obtain the following relation: For a given arbitrary pair of states represented by $(\Theta_\alpha,\Theta_\beta)$ and $(\Theta_{\alpha'},\Theta_{\beta'})$, the following relation holds:

$$(\Theta_\alpha,\Theta_\beta)\prec(\Theta_{\alpha'},\Theta_{\beta'}) \text{ if and only if } S(\Theta_\alpha)+S(\Theta_\beta)\le S(\Theta_{\alpha'})+S(\Theta_{\beta'}), \tag{24}$$

where Θ_j, $\mathrm{j}=\alpha,\beta,\cdots$, is the element of the state space $\mathcal{M}_\Theta\times\mathcal{M}_\Theta$. It should be noted that each Θ can depend on different values of *L*. This is the entropy principle expressed in terms of Θ's for a quantum compound system. The statement (24) means that $\left(\Theta_{\alpha'},\Theta_{\beta'}\right)$ is adiabatically accessible from $\left(\Theta_{\alpha},\Theta_{\beta}\right)$. We note that this statement is useful when one considers the heat transfer between quantum systems, *e.g.*, thermal contact. We will discuss thermal contact in a separate paper [12].

Now we consider a relation between parameter Θ and thermodynamic temperature T. We start with introducing thermodynamic temperature T that is defined as a partial derivative of entropy S with respect to U:

$$\frac{1}{T}:=\left.\frac{\partial S\left(U,V\right)}{\partial U}\right|_{V}, \tag{25}$$

where U and V denote the internal energy and volume (work coordinate) of a system, respectively [2] [3]. In the definition (25), differentiability of the entropy function S may be ensured by a concavity of the entropy function.[5]

Here, we introduce a corollary **C3**:

C3. There exists a one-to-one correspondence between T and Θ.

Proof of C3. From the definition of thermodynamic temperature (25), the value of T is defined for the respective values of U owing to a uniqueness of the differential coefficient of a concave function S. Let E be a value of internal energy U $\left(=E\right)$. By **C1**, the value E has one-to-one correspondence with the value of Θ. Therefore, thermodynamic temperature T can be determined by the parameter Θ in $\mathcal{M}_{\Theta}$ introduced in Sec. 3. □

This corollary establishes the consistency between the parameter Θ and thermodynamic temperature T. Let us consider the partial derivative in the right-hand side of Equation (25). We note that a value of T is determined by Θ and L. This implies that an internal energy plays the same role as temperature in quantum system (see **C1**). This agrees with the proposition of Bender *et al.* [9]. Equation (25) is useful to obtain the relation between thermodynamic temperature and quantum states characterized by Θ. This is also useful to obtain a structure in the thermodynamic state space $\mathcal{M}_{\Theta}$.

In this section, we have established the representation of the second law in terms of the entropy S for quantum systems as a function of parameter Θ, so that one can define *thermodynamic temperature* T as a function of a state parameter Θ: $T=T\left(\Theta\right)$.

3.4. The Third Law

Let us briefly discuss the third law of thermodynamics for quantum systems. The third law requires quantum states [1]. Therefore, it must take into account the probabilistic nature of quantum theory. The probability function $P\left(\Theta\right)$ is

---

[5] A relation between thermodynamic temperature and concavity of the entropy function is discussed by Lieb and Yngvason [3].

found from probability amplitudes (see **L1** in Sec. 3). Accordingly we can obtain the representation of the third law: The entropy S is equal to zero *only* when the probability function satisfies the condition:

$$P(\Theta; x=i)=1 \text{ for arbitrary } i. \tag{26}$$

At the absolute zero temperature, one can expect a state of quantum system being in a *single* state such as

$$|\Psi\rangle = a_i |i\rangle \text{ for arbitrary } i. \tag{27}$$

The single state here means that only one outcome is obtained with a probability one by maximal tests.

4. Results and Discussion

In this paper, we presented a representation of the basic laws in thermodynamics for quantum system in a pure state and investigated a relationship between thermodynamics and quantum theory. We obtained the representation of the zeroth, first, second and third laws in quantum thermodynamics, which affords the key to treat quantum system thermodynamically. In the derivation of the zeroth law that assures the existence of a property called "temperature" in classical thermodynamics, we proposed parameters Θ that characterize the equivalence law among state vectors, leading to the zeroth law of quantum thermodynamics. We showed that the existence of such parameters depends on the probabilistic nature of quantum theory (**L1**). It should be however noted that the parameters Θ's introduced in our formulation make sense only when there exists the map $\mathcal{M}_\Psi \mapsto \mathcal{M}_\Theta$, *i.e.*, $\mathcal{M}_\Theta$ is an injection of $\mathcal{M}_\Psi$. Therefore, in order to establish the representation of the zeroth law in terms of the elements in $\mathcal{M}_\Theta$, *same* maximal tests must be performed to obtain the corresponding elements of $\mathcal{M}_\Theta$ from the respective elements in $\mathcal{M}_\Psi$. Though this statement can refer only to the case where the work coordinates are fixed, the second law releases this restriction.

We established the representation of the first law, where the internal energy of quantum system as a state quantity is specified by the parameters in $\mathcal{M}_{\Theta,L}$ (**C1**). From the first law, when the state changes from Θ_α to Θ_β, heat entering the system is represented by $d'Q = U(\Theta_\beta) - U(\Theta_\alpha)$ in the case where $d'W = 0$ (*i.e.*, the work coordinate is fixed). Thus one can measure the quantity of heat in terms of Θ's for quantum systems. By considering the adiabatic process ($d'Q = 0$), one can also quantify the work done to the system in terms of Θ's in the present theory. Since the term $d'W$ is due to a shift of the outcomes caused by the variations of work coordinate, one can measure the work $d'W$ in terms of the changes in the internal energy as $d'W = U(\Theta_\alpha, L_1) - U(\Theta_\alpha, L)$ when we consider an adiabatic process, where Θ_α remains constant. Here $U(\Theta_\alpha, L_1)$ describes the internal energy of the final state. This is useful to obtain the relation among $\{\mathcal{M}_{\Theta,L_1}, \mathcal{M}_{\Theta,L_2}, \cdots\}$ and the relation gives a structure in thermodynamic state space $\mathcal{M}_{\Theta,L}$.

We also obtained the representation of the second law of thermodynamics for quantum system. The representation depends on **L2** and assures the entropy principle. The representation of the second law (entropy principle) proposed here would give some insight into the structure (*order relation*) in $\mathcal{M}_{\Theta,L}$ for quantum systems. Therefore, this would afford one to investigate what thermodynamic structure exists in a state space describing thermodynamics of quantum systems in terms of those parameters in $\mathcal{M}_{\Theta,L}$.

The third law requires as a principle that entropy should be zero at the absolute zero temperature [1]. In our representation, entropy can take the value of zero when the state is described by $\left|\Psi^{\mathrm{j}}\right\rangle = b_{\mathrm{ij}}\left|i\right\rangle$ (*i.e.*, the state of the system is characterized by a *single* state for each ensemble j). Since one can see the existence of a unique function of Θ that describes the thermodynamic temperature, there is *no* necessity of referring to an explicit form of the function. Let $\left|i\right\rangle$ be an energy eigenstate of the system. At the absolute zero temperature, the state is occupied from the lower states and the state $\left|\Psi^{\mathrm{j}}\right\rangle = b_{\mathrm{ij}}\left|i\right\rangle$ should correspond to the lowest energy level *i*. Therefore, we required one more condition to the statement Equation (26), that is, the *only* state *i* is reserved for the description of the lowest energy level. It should be noted that this requirement is consistent with Fermi and Bose statistics at the absolute zero temperature and the state is free from the type of Hamiltonian for quantum system, establishing the third law of quantum thermodynamics. By the third law along with the definition of entropy function (22), one can determine the absolute value of entropy.

Finally we consider the problem of thermal contact. This problem makes the notion of heat transfer much clear. Let us consider a compound system comprised of subsystem A and B whose initial states are characterized by ($\Theta_\alpha, \Theta_\beta$), respectively. Without loss of generality we could use the same value of the work coordinate for both systems. The compound system as itself is isolated from any other environment. After thermal contact the state of each subsystem becomes ($\Theta_\gamma, \Theta_\gamma$) due to the heat transfer between subsystems. For such a case, one can predict a possible value of Θ_γ. Since the case satisfies the inequality (24), we could evaluate the values of Θ by obtaining the probability function $P(\Theta; x)$. Therefore one can verify the order relation among Θ's (*i.e.*, whether $(\Theta_\alpha, \Theta_\beta) \prec (\Theta_\gamma, \Theta_\gamma)$ is true or not). In other words, by studying the order relation among parameters Θ's through the study of respective entropies, thermodynamic structure hidden in quantum systems could be elucidated.

Now we discuss again the validity of the zeroth law (8) from a viewpoint of the thermal contact. If the subsystem A and B are in thermal equilibrium, then there is no heat transfer between them. Thus one can say Θ's are the parameters which characterize respective thermal equilibrium states of subsystems A and B.

By applying the quantum thermodynamics presented in this paper, thermodynamic nature of quantum systems could be studied in all temperature without considering detailed mechanics of quantum systems. This is the advantage to utilize quantum thermodynamics presented in this paper. We hope that the present theory sheds light on further understanding a relationship between thermodynamics and quantum theory.

Conflicts of Interest

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

References

[1] Fermi, E. (1956) Thermodynamics. Dover, New York.

[2] Callen, H.B. (1985) Thermodynamics and an Introduction to Thermostatics. 2nd Edition, Wiley, New York.

[3] Lieb, E.H. and Yngvason, J. (1999) The Physics and Mathematics of the Second Law of Thermodynamics. *Physics Reports*, **310**, 1-96. https://doi.org/10.1016/S0370-1573(98)00082-9

[4] von Neumann, J.L. (1932) Mathematische Grundlagen der Quantenmechanik. Springer, Berlin. [Beyer, E.T. (Trans.) (1955) Mathematical Foundations of Quantum Mechanics. Princeton University Press, Princeton.]

[5] Peres, A. (1993) Quantum Theory: Concepts and Methods. Kluwer-Academic, Boston.

[6] Wehrl, A. (1978) General Properties of Entropy. *Reviews of Modern Physics*, **50**, 221. https://doi.org/10.1103/RevModPhys.50.221

[7] Partovi, M.H. (1989) Quantum Thermodynamics. *Physics Letters A*, **137**, 440-444. https://doi.org/10.1016/0375-9601(89)90221-1

[8] Scully, M.O., Zubairy, M.S., Agarwal, B.S. and Walther, H. (2003) Extracting Work from a Single Heat Bath via Vanishing Quantum Coherence. *Science*, **299**, 862-864. https://doi.org/10.1126/science.1078955

[9] Bender, C.M., Brody, D.C. and Meister, B.K. (2000) Quantum Mechanical Carnot Engine. *Journal of Physics A*: *Mathematical and General*, **33**, 4427-4436. https://doi.org/10.1088/0305-4470/33/24/302

[10] Isham, C.J. (1995) Lectures on Quantum Theory: Mathematical and Structural Foundations. Imperial College Press, London. https://doi.org/10.1142/p001

[11] Born, M. and Fock, V. (1928) Beweis des Adiabatensatzes. *Zeitschrift für Physik*, **51**, 165-180. https://doi.org/10.1007/BF01343193

[12] Taira, H. and Suzuki, A. Representaion of Thermodynamic Laws in Quantum Systems II. (To be submitted)

Journal of Modern Physics, 2018, 9, 2437-2442
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Carrier-Induced Magnetic Solitons and Metal-Insulator Transition in Diluted Magnetic Semiconductors $Ga_{1-x}Mn_xAs$

I. Kanazawa[1], S. Nakamura[1], R. Maeda[2]

[1]Department of Physics, Tokyo Gakugei University, Tokyo, Japan
[2]Institude of Multidisciplinary Research for Advanced Materials (IMRAM), Tohoku University, Sendai, Japan
Email: kanazawa@u-gakugei.ac.jp, ryota.maeda.p4@dc.tohoku.ac.jp

How to cite this paper: Kanazawa, I., Nakamura, S. and Maeda, R. (2018) Carrier-Induced Magnetic Solitons and Metal-Insulator Transition in Diluted Magnetic Semiconductors $Ga_{1-x}Mn_xAs$. *Journal of Modern Physics*, **9**, 2437-2442.
https://doi.org/10.4236/jmp.2018.914156

Received: September 7, 2018
Accepted: December 11, 2018
Published: December 14, 2018

Abstract

We discuss hole-induced magnetic solitons and metal-insulator transition of transport properties in diluted magnetic semiconductors $Ga_{1-x}Mn_xAs$ from the standpoint of a field theoretical formulation, and analyze experimental data of transport properties, using the supersymmetry sigma formula and the effective Lagrangian of diffusion model.

Keywords

Diluted Magnetic Semiconductor, Magnetic Soliton, Metal-Insulator Transition, Localization

1. Introduction

Diluted magnetic semiconductors (DMSs), which are formed by substitution of several percent of cation sites in a host semiconductor with magnetic impurities, are actively investigated both theoretically and experimentally, due to their potential applications in new generations of semiconductor spintronic devices [1]. Because the carriers in DMSs are considered to mediate the magnetic interaction between the magnetic ions [2], these materials are very important for semiconductor-based spintronic devices to control the spin degree of freedom of the carriers. Due to the mediation mechanism, the ferromagnetism in DMSs is called carrier-induced ferromagnetism. Prototypical DMS systems such as $Ga_{1-x}Mn_xAs$ and $In_{1-x}Mn_xAs$ show severely limited chemical solubility due to the substitution of divalent Mn atoms for the trivalent Ga or In sites. In order to prevent phase separation, these materials should be grown at low temperature (T from 200˚C to 300˚C), which results in an abundance of different types of crystal defects. As a result, a theoretical study of DMSs is very difficult owing to two factors (strong disorder and exchange interaction), which must be taken into account nonperturbatively.

Understanding the mechanism behind the carrier-induced ferromagnetism is of significance for further development of semiconductor spintronic devices. Several theoretical models for carrier-induced ferromagnetism in (Ga, Mn)As have been proposed [1]-[6]. In addition, interesting phenomena such as the photo-induced magnetic polaron in DMSs have been reported [7] [8] [9] [10]. These studies stimulate us to investigate the hole-induced magnetic solitons. It has been required to consider the behavior of the hedgehog-like magnetic soliton and the domain wall from a viewpoint of quantum theory. Kanazawa [11] has discussed the hole-induced magnetic solitons in DMSs from the standpoint of a field-theoretical formulation. Metal-insulator transition (MIT) and large magnetoresistance (MR) effects in DMSs (Ga, Mn)As have been reported [12] [13] [14] [15] [16]. Kanazawa and coworkers [17] [18] [19] [20] [21] have discussed these anomalous properties in DMSs theoretically.

In this study, the anomalous transport properties in DMSs are discussed using a field-theoretical formulation. Then we analyze some conductivity data in DMSs (Ga, Mn)As, using the gauge-invariant effective Lagrangian density and quantized magnetic solitons.

2. A model System and Hole-Induced Magnetic Solitons

According to the aggregation of hole-induced magnetic solitons, the non-monotonic temperature dependence of the transport properties of (Ga, Mn)As is qualitatively explained as being due to the hole localization around the Mn ions. It has been suggested that the ferromagnetic ordering might be due to a double-exchange-like interaction and the remarkable change of spin exchange interaction among Mn ions by the hole seems to be cooperative and non-linear (Yang Mills like). Kanazawa and coworkers [22] [23] [24] [25] have proposed that in quasi-(2 + 1) dimensions in a quantum antiferromagnet the hole-induced magnetic disorder leads to hedgehog-like solitons, which are composed of the doped hole and the cloud of SU(2) Yang-Mills fields with spin disorder around the hole. In addition, based on the important ideas in Refs. [26] [27] [28] [29], it has been proposed that the hedgehog-like soliton in a three-dimensional system is specified by rigid-body rotation, which is related to gauge fields of SO(4) symmetry for $\mathcal{S}^3$ [30] [31] [32] [33] [34].

Then the Yang-Mills fields A_μ^a induced by the doped hole have a local SO(4) symmetry. Here we have thought that the SO(4) symmetry fields A_μ^a are spontaneously broken around the hole through the Anderson-Higgs mechanism, in the III-V-based diluted magnetic semiconductors with magnetic manganese ion-doping. Through the spontaneous symmetry breaking $\langle 0|\phi_a|0\rangle = \langle 0,0,0,\mu\rangle$, the effective Lagrangian density has been introduced [11] [19]. That is, the effective Lagrangian density reveals that the ferromagnetically aligned Mn spins create the cluster, in which the hole is trapped, with the radius $R_c \sim 1/m_2$. Katsumoto *et al.* [16] have shown that the localization length l_c of the wave function of holes plays an important role in the metal-insulator transition in DMS (Ga, Mn)As. It is suggested strongly that the l_c might correspond to $R_c \sim 1/m_2$. In III-V-based DMSs, the resistivity increases remarkably as the temperature decreases. In addition, in the same temperature region, the negative magnetoresistance grows rapidly as the temperature decreases. To explain the electron hopping and spin dynamics, we introduce an effective Hamiltonian, *H*, for the magnetic soliton $O\left(r_{\tilde{i}}\right)$ [18] [20]

$$H = -J\sum_{\langle \tilde{i},\tilde{j}\rangle}\cos\left(\theta_{\tilde{i}\tilde{j}}/2\right)O\left(r_{\tilde{i}}\right)\cdot O\left(r_{\tilde{j}}\right)+\frac{1}{2}K\sum_{\tilde{i}\neq\tilde{j}}\frac{O\left(r_{\tilde{i}}\right)\cdot O\left(r_{\tilde{j}}\right)}{\left|r_{\tilde{i}}-r_{\tilde{j}}\right|}. \tag{1}$$

Here the first sum $\sum_{\langle\tilde{i},\tilde{j}\rangle}$ is taken only over nearest neighbors (the distance between each magnetic soliton is $\leq 2R_c$), while the second sum is taken over all pair($\tilde{i}\neq\tilde{j}$ means $\left|r_{\tilde{i}}-r_{\tilde{j}}\right| > 2R_c$) [18] [20]. $\theta_{\tilde{i}\tilde{j}}$ is the angle between $N_{\tilde{i}}$ and $N_{\tilde{j}}$. Here $N_{\tilde{i}}$ and $N_{\tilde{j}}$ represent the effective spins of the solitons $O\left(r_{\tilde{i}}\right)$ and $O\left(r_{\tilde{j}}\right)$, respectively. $N_{\tilde{i}}$ is the summation of the ferromagnetic spin, N_i, of Mn within $\sim(4/3)\pi R_c^3\left(\tilde{i}\right)$ around the hole at the site $r_{\tilde{i}}$. *J* is defined in Equation (2). *K*, which is introduced in Ref. [18], is the effective long-range interaction constant. The first term shows short-range ferromagnetic ordering interaction and the second one shows long-range frustration.

$$-J = -\frac{g_2^2}{4\pi}\frac{e^{-m_1 r}}{r}\bigg|_{r\sim\frac{1}{m_1}} \sim -\frac{g_2^2}{4\pi}m_1 e^{-1} \tag{2}$$

is the short-range attractive potential, which is derived from massive gauge fields A_μ^1, A_μ^2, and A_μ^3 exchange interaction. When the magnetic soliton, $O\left(r_{\tilde{i}}\right)$, with the effective spin $N_{\tilde{i}}$ is located at the nearest-neighbor site of the magnetic soliton, $O\left(r_{\tilde{j}}\right)$, with the effective spin $N_{\tilde{j}}$, holes are hopping between the two solitons $O\left(r_{\tilde{i}}\right)$ and $O\left(r_{\tilde{j}}\right)$. If $N_{\tilde{i}}$ is parallel to $N_{\tilde{j}}$, the p-d exchange interaction induces large reduction of the kinetic energy. The hopping term between the nearest neighbors of hedgehog-like solitons (clusters) leads to an additional term in the σ-model describing a coupling of the supermatrices, $Q_{\tilde{i}}$, corresponding to different magnetic solitons (clusters) [18]. We discuss the transport properties of DMSs for connected clusters, where the radius is $R_c \sim 1/m_2$, of DMSs. Approximately we introduce the following approximate free energy by using the formula for the model of granulated clusters [35] [36]

$$\tilde{F}(Q) = str\left(-\sum_{\langle\tilde{i},\tilde{j}\rangle}J_{\tilde{i}\tilde{j}}Q_{\tilde{i}}Q_{\tilde{j}}+\frac{i}{4}(\omega+i\delta)\sum_{\tilde{i}}\Delta_{\tilde{i}}^{-1}\Lambda Q\right), \tag{3}$$

where $J_{\tilde{i}\tilde{j}} = J\cos\left(\theta_{\tilde{i}\tilde{j}}/2\right)\frac{1}{\Delta_{\tilde{i}}\Delta_{\tilde{j}}}$. Then $\Delta_{\tilde{i}}$ is the mean energy level spacing at the hedgehog-like soliton (cluster) $O\left(r_{\tilde{i}}\right)$ and $J > 0$. The diffusion coefficient D_0 is introduced as follows,

$$
\begin{aligned}
D_0 &\sim \frac{4\Delta}{\pi}\sum_{\tilde{j}} J_{\tilde{i}\tilde{j}}\left(r_{\tilde{i}}-r_{\tilde{j}}\right)^2 \\
&\sim \frac{4\Delta}{\pi}\sum_{\tilde{j}} J\cos\left(\theta_{\tilde{i}\tilde{j}}/2\right)\frac{1}{\Delta^2}\left(r_{\tilde{i}}-r_{\tilde{j}}\right)^2 \\
&\sim \frac{4}{\pi\Delta}\sum_{\tilde{j}} J\cos\left(\theta_{\tilde{i}\tilde{j}}/2\right)\left(r_{\tilde{i}}-r_{\tilde{j}}\right)^2
\end{aligned}
\tag{4}
$$

Here $\Delta = \frac{1}{\nu\pi R_c^2}$ and ν is the density of states of the carriers at the Fermi surface. In the case of the low frequency limit of ω, the localization length L_{loc} is shown as follows,

$$L_{loc} \propto \pi^2 \nu R_c^2 D_0 \sim \pi^2 \nu R_c^2 \frac{4}{\pi\Delta}\sum_{\tilde{j}} J\cos\left(\theta_{\tilde{i}\tilde{j}}/2\right)\left(r_{\tilde{i}}-r_{\tilde{j}}\right)^2 \tag{5}$$

We shall consider the variable range hopping conductivity and the system length $L \gg L_{loc}$ as follows,

$$\sigma \propto \exp\left[-\left(A/T\right)^{1/(d+1)}\right] \tag{6}$$

where d is the dimensionality of the system.

$$A \propto \left(\frac{1}{L_{loc}}\right)^{d/(d+1)} \sim \left(\frac{1}{\pi^2 \nu R_c^2 D_0}\right)^{d/(d+1)} \tag{7}$$

Figure 1 shows the temperature dependence of the conductivity σ for as-grown and annealed samples (experimented data) [15] and the fitting lines (solid lines) of $Ga_{0.95}Mn_{0.05}As$. The annealing is performed at 310˚C for 15 mins.

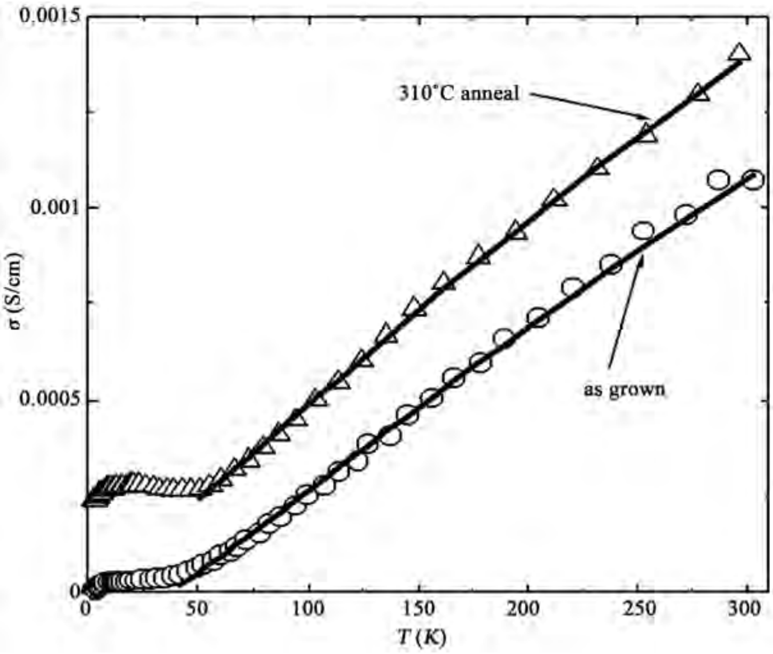


Figure 1. Temperature dependence of conductivity of the diluted semiconductor $Ga_{0.95}Mn_{0.05}As$ [15]. The anniling time is 15 min. The solid lines have been fitted to the measured data.

The as-grown sample and the sample annealed at 310˚C show insulating behavior above ~30 K and ~50 K, respectively. The annealing at 310˚C increases the conductivity. Annealing might reduce concentration of As antisites and interstitial Mn. As the conductivity σ increases, the high-temperature structure moves to higher temperatures, which means T_c (Curie temperature) increases. Thus the concentration ρ of mobile holes and T_c are enhanced by the annealing. The experimental data are fitted well with Equations (6) and (7), as shown with solid lines in Figure 1. Comparing the experimental data (annealing at 310˚C) with those (as grown), it is thought that the value of L_{loc} (after annaealing at 310˚C) is much larger than of those (as grown), as seen from Equation (7).

3. Conclusion

The hole-induced magnetic solitons and metal-insulating transition of transport properties in DMSs have been discussed based on a field theoretical formulation. We have analyzed experimental data on the transport properties of GaMnAs by using the effective Lagrangian of diffusion model.

Conflicts of Interest

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

References

[1] Dietl, T. (2010) *Nature Materials*, **9**, 965-974. https://doi.org/10.1038/nmat2898

[2] Dietl, T. and Ohno, H. (2014) *Reviews of Modern Physics*, **86**, 187. https://doi.org/10.1103/RevModPhys.86.187

[3] Akai, H. (1998) *Physical Review Letters*, **81**, 3002. https://doi.org/10.1103/PhysRevLett.81.3002

[4] König, J., Lin, H.-H. and MacDonald, A.H. (2000) *Physical Review Letters*, **84**, 5628.

[5] Berciu, M. and Bhatt, R.N. (2001) *Physical Review Letters*, **87**, 107209. https://doi.org/10.1103/PhysRevLett.87.107203

[6] Kaminski, A. and Dag Sarma, S. (2002) *Physical Review Letters*, **88**, 247202. https://doi.org/10.1103/PhysRevLett.88.247202

[7] Chapman, R.A. and Hutchinson, W.G. (1967) *Physical Review Letters*, **18**, 443. https://doi.org/10.1103/PhysRevLett.18.443

[8] Emin, D. (1993) *Physical Review B,* **48**, 13691. https://doi.org/10.1103/PhysRevB.48.13691

[9] Koshihara, S., *et al.* (1997) *Physical Review Letters*, **78**, 4617. https://doi.org/10.1103/PhysRevLett.78.4617

[10] Hirakawa, C.K., *et al.* (2002) *Physical Review B*, **65**, 193312. https://doi.org/10.1103/PhysRevB.65.193312

[11] Kanazawa, I. (2002) *Phase Transitions*, **45**, 1043-1047. https://doi.org/10.1080/01411590290034218

[12] Oiwa, A., *et al.* (1997) *Solid State Communications*, **104**, 209-213. https://doi.org/10.1016/S0038-1098(97)00178-6

[13] Oiwa, A., *et al.* (1998) *Physica B*, **246-251**, 775.

[14] Iye, Y., *et al.* (1998) *Materials Science and Engineering: B*, **63**, 88-95. https://doi.org/10.1016/S0921-5107(99)00057-4

[15] Hayashi, T., *et al.* (2001) *Physica E*, **10**, 130-134. https://doi.org/10.1016/S1386-9477(01)00067-4

[16] Katsumoto, S., *et al.* (2001) *Materials Science and Engineering: B*, **84**, 88. https://doi.org/10.1016/S0921-5107(01)00575-X

[17] Kanazawa, I. (2004) *Physica E*, **21**, 961. https://doi.org/10.1016/j.physe.2003.11.171

[18] Kanazawa, I. (2005) *Physica E*, **29**, 647. https://doi.org/10.1016/j.physe.2005.06.046

[19] Kanazawa, I. (2006) *Physics Letters A*, **255**, 460. https://doi.org/10.1016/j.physleta.2006.02.040

[20] Kanazawa, I. (2007) *Physica E*, **40**, 277. https://doi.org/10.1016/j.physe.2007.06.012

[21] Kanazawa, I. (2008) *Physica Status Solidi* (*c*), **5**, 318. https://doi.org/10.1002/pssc.200776594

[22] Kanazawa, I. (2003) *Journal of Physics A*, **36**, 9371. https://doi.org/10.1088/0305-4470/36/35/321

[23] Kanazawa, I. and Maeda, R. (2017) *Journal of Superconductivity and Novel Magnetism*, **30**, 49. https://doi.org/10.1007/s10948-016-3853-z

[24] Kanazawa, I. and Maeda, R. (2018) *Journal of Superconductivity and Novel Magnetism*, **31**, 671. https://doi.org/10.1007/s10948-017-4363-3

[25] Kanazawa, I. and Maeda, R. (2018) *Quantum Studies: Mathematics and Foundations*, **5**, 141.

[26] Kleman, M. and Sadoc, T.F. (1979) *Journal of Physics*, **40**, 567.

[27] Kanazawa, I. (1992) *Journal of Non-Crystalline Solids*, **150**, 271. https://doi.org/10.1016/0022-3093(92)90135-7

[28] Kanazawa, I. (1993) *Physics Letters A*, **176**, 246. https://doi.org/10.1016/0375-9601(93)91043-5

[29] Kanazawa, I. (1997) *Progress of Theoretical Physics Supplements*, **126**, 397. https://doi.org/10.1143/PTPS.126.393

[30] Nelson, D.R. (1983) *Physical Review Letters*, **50**, 982. https://doi.org/10.1103/PhysRevLett.50.982

[31] Sethna, J.P. (1983) *Physical Review Letters*, **51**, 2198. https://doi.org/10.1103/PhysRevLett.51.2198

[32] Sachder, S. and Nelson, D.R. (1984) *Physical Review Letters*, **53**, 1947. https://doi.org/10.1103/PhysRevLett.53.1947

[33] Kanazawa, I. (2001) *Journal of Non-Crystalline Solids*, **293-295**, 615. https://doi.org/10.1016/S0022-3093(01)00759-1

[34] Kanazawa, I. (2002) *Journal of Non-Crystalline Solids*, **312-314**, 608. https://doi.org/10.1016/S0022-3093(02)01765-9

[35] Efetov, K.B. (1983) *Advances in Physics*, **32**, 53. https://doi.org/10.1080/00018738300101531

[36] Efetov, K.B. (1992) *Physical Review B*, **45**, 11546. https://doi.org/10.1103/PhysRevB.45.11546

Journal of Modern Physics, 2018, 9, 2443-2456
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

A Photonic Model of the Big Bang

J. G. Lartigue*

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

How to cite this paper: Lartigue, J.G. (2018) A Photonic Model of the Big Bang. *Journal of Modern Physics*, **9**, 2443-2456.
https://doi.org/10.4236/jmp.2018.914157

Received: November 14, 2018
Accepted: December 17, 2018
Published: December 20, 2018

Abstract

There are two main theories about the origin of the Universe that show similitude with the Genesis writings, though in different verses: the Big Bang[1] and the eternal Universe[2] (an eventual quantum fluctuation). However, it is possible to partially include the quantum theory in the Big Bang thanks to the nature of photons, to obtain a simple model. It is assumed as the origin of the Universe (space, time, matter and physical laws). A subsequent enormous expansion has been explained by a supposed brief Inflation period, followed up today by a constant adiabatic expansion acceleration. This paper assumes that the Universe is the total Space which contains the Physical Universe covered by an external, empty Space, both expanding at a constant Hubble acceleration Γ_H [1]. A Big Bang design is intended by a deduction of the energy and number of primeval photons, from the present CMB value; they would have reacted whether to generate the Physical Universe or to decay till the CMB level. It follows an approach to the Universe expansion work, based on the Hubble field (V_H) as well as on Thermo-dynamics. They are calculated: the time and angular momentum required for the Physical Universe to reach the maximum internal velocity c as well as, simultaneously, a c tangential velocity. The Universe luminosity at different periods and the adequate expressions of parameters (Ω, q, k) are revised. It is proposed a modification in the equation of the $H(t)$ parameter and the H_o value. The operator of convective derivative is applied to obtain an equation of continuity for the photonic energy; an adiabatic Jacobian gives similar results. This essay differs from others based on black box radiation, since the Universe has no walls and the photons energy decays continuously.

Keywords

Hubble Parameters, Space Acceleration, Photonic Equation of Continuity, Universe Angular Momentum

---

*Retired Professor.

[1] Genesis, Chapter 1, Verse 3: And God said "let there be light" and there was light.

[2] Genesis, Chapter 1, Verse 2: "The earth was without form and void… and the spirit of God was hovering over the face of waters".

1. Introduction

1.1. Previous Models

Reference [2] describes the extensive research developed to determine the CRB variations, which may suggest a spectral distribution of the Big Bang original photons. Reference [3] is a complete text on Statistical Physics though its chapter on photon gas is mainly oriented to black body and cavities radiations. A previous classical book [4] mentions the Wien constant, necessary to determine the maximum energy density, though it is exemplified only by solar radiation. Reference [5] assumes that the photon gas is a carrier for electromagnetic waves in a very complex model. However, it was not found a Big Bang thermo-dynamical model that would, probably, exist.

1.2. The Selected Universe Model

The Einstein gravitational equation, including the cosmological constant Λ, was modified by De Sitter for an empty Universe [6]. In reference [7] the $\Lambda g_{\mu\nu}$ tensor was substituted by a Hubble tensor (whose scalar equals $3H^2$); it implies a Hubble positive potential field $V(r)$ and the expansion acceleration Γ_H. The photonic model has been selected because the Planck temperature represents an initial energy about 10^{22} (Mev), which overpasses the binding energy of any mass, from neutrino to Higgs particle. For simplicity it is initially assumed as a mono-energetic photonic source, which could later be modified as function of a probable Planck Length indetermination.

1.3. The Light Speed in Dense Matter

A doubt could be elicited due to the use of the **c** value in Planck parameters because, at that time, the linear speed of initial photons could not be higher than that of the space expansion. However, such an obstacle may have overcome by the results of a research on opaque bodies [8] which concluded that the total trajectory of light through a big number of collisions per second corresponds to ***c***. That would confirm the light velocity as a true constant in the Physical Universe, provided that the internal mean free path in the mass would be wide enough to admit a wavelength displacement; obviously, the ***c*** value had been confirmed in vacuum.

2. Some Universe Parameters

2.1. Previously Published Parameters

The maximum velocity of matter into the physical Universe is ***c*** (**Figure 1**); the present time is $t_o = 14$ Gy. It has been assumed that the Big Bang duration was $t_{be} = 10^{12}$ (s) [9] or 3.2×10^4 (y). Besides, the maximum velocity of matter into the physical Universe has been limited to ***c***, which was freely manifest at the Universe expansion time $t_c = 1.1 \times 10^{17}$ (s); the corresponding Universe scale factor R_c (Universe radius $\boldsymbol{r}_c$), where velocity expansion **c** freely occurred, was $\boldsymbol{r}_c = 1.7 \times 10^{27}$ (cm). Otherwise, it has been shown [1] that the space acceleration

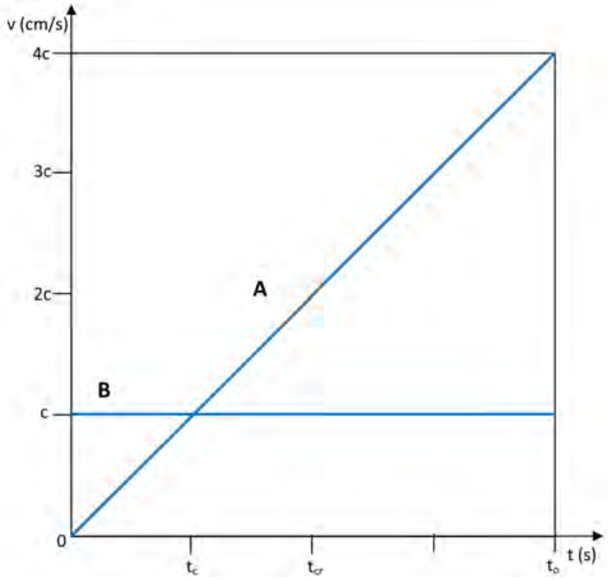


Figure 1. Universe expansion velocity as a time function. A: expansion velocity of Space, Equation (1). B: constant velocity of light in the Physical Universe.

$\mathbf{\Gamma}_H$ (cm·s^{-2}) is a constant, independent of mass presence. Present time is t_o. The Space expansion velocity is:

$$\mathbf{v}_H = \mathbf{\Gamma}_H \cdot t \tag{1}$$

The value of the Hubble parameter has been here corrected to $H_o = 3.2 \times 10^{-18}$ (s^{-1}) assuming a maximum speed of 100 (km/sMpc) in the definition equation. So, the intensity of the Hubble field is $\mathbf{\Gamma}_H = 2.65 \times 10^{-7}$ (cm·sec^{-2}) (**Figure 2**).

2.2. Additional Parameters

1) The critical condition of the Universe has been defined in two ways: respect to a ratio Ω [10] of present and critical densities and respect to a net acceleration [11] as:

$$\Omega = \rho_o / \rho_{cr} \tag{2}$$

$$\mathbf{\Gamma}_N = \mathbf{\Gamma}_H - \mathbf{\Gamma}_G \tag{3}$$

$$\mathbf{\Gamma}_H = H^2 \mathbf{r} = \text{constant} \tag{4}$$

$$\mathbf{\Gamma}_G = GM / \mathbf{r}^2 \tag{5}$$

Though Equations (4) and (5) may be here applied, Equations (2) and (3) should not because of the following: Equation (2) is not adequate for the De Sitter model since the vacuum density of the external empty space is $\rho_o = 0$ and, therefore, $\Omega = 0$, independently of the density of the physical Universe. A critical point could rather correspond to the t_c time, when the expansion velocity of Space and the internal velocity in the Physical Universe diverge; at that time, the gravitational acceleration is $\Gamma_G \sim 10^{-2}$ times lower than the space acceleration Γ_H (**Figure 2**). Even more, the eventual numerical equality in Equation (3) occurs at a time $t_{cr} \sim 2t_o$, without any physical consequence.

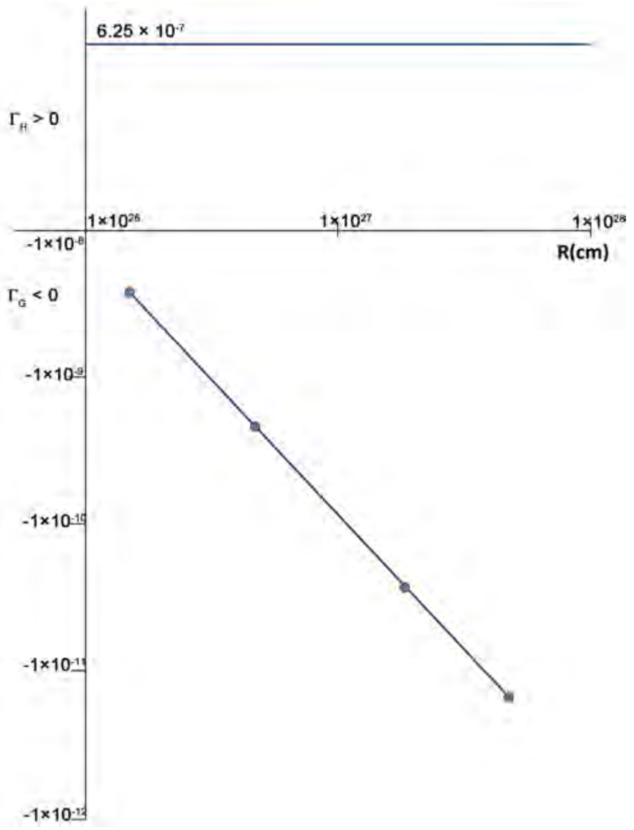


Figure 2. Evolution of the Hubble (Γ_H) and gravitational (Γ_G) acceleration (cm·s^{-2}) as function of the radial factor (R), Equations (4) and (5).

2) Another important factor is the so-called deceleration factor q_u defined [12] by the equation:

$$q_u = -R\ddot{R}/\dot{R} \tag{6}$$

This expression was considered adequate at the time when the Universe acceleration was assumed negative, accordingly to the second Friedman equation [13]:

$$\frac{\ddot{R}}{R} = -\frac{4\Pi G}{3}\left(\rho + \frac{3p}{c^2}\right) \tag{7}$$

In these two equations R is the scale factor of the Universe, $\dot{R}$ is the expansion velocity and $\ddot{R}$ is the acceleration; as well, ρ is the Universe density and p is its pressure.

To date, at a positive acceleration, the minus sign should be discarded from Equation (6).

3) The curvature of an spherical Universe may be estimated by the classical expression $k = 1/(R_u)^2$. If obtained from the first Friedman equation, this gives the following results: at early times, when the Hubble parameter was higher than today, k would be negative, *i.e.* a hyperbolic geometry; now, at the smaller magnitude of H_o, the k value is still negative though in a rank of 10^{-40} (cm^{-2}) that suggest a flat Universe. Otherwise, the FLRW equation admits 3 possible values

for $k\,(+1, 0, -1)$ to cover all of curvature possibilities.

4) The Laplacian for the Hubble potential was deduced [1] as $\nabla^2 V_H = 0$, since V_H is not a function of the mass presence.

3. The Hubble Parameters

1) The original parameter was defined by *H.* Hubble as the relation $H = v/c$; later, it has been defined by its own units (s^{-1}) in successive concepts: from a universal constant (H_o), to a single reciprocal function of time, $H = (1/t)$ and even to $H = 2/3t$ in the Einstein-De Sitter model.

2) The Hubble time has been defined as $t_H = 1/H$. However these equations have been applied for a Universe whose expansion velocity was assumed constant. If the acceleration ($\boldsymbol{\Gamma}_H$) is the true constant, the Hubble time should be expressed as:

$$t_H = \sqrt{2}/H(t)\,(\text{s}) \tag{8}$$

3) The Hubble Length, as known nowadays, is defined by the equation:

$$L_H = c/H_o\,(\text{cm}) \tag{9}$$

Substitution of Equation (8), gives:

$$L_H = ct_o/\sqrt{2}\,(\text{cm}) \tag{10}$$

Equation (9) is equivalent to that for an inertial frame: $\boldsymbol{r} = \boldsymbol{v}t$, if $\boldsymbol{v}$ had the $\boldsymbol{c}$ value. Equation (10) was obtained for an accelerated frame, $\boldsymbol{\Gamma}_H$. So, this equation would be valid for the case when the expansion velocity would be higher than $\boldsymbol{c}$ as: $\boldsymbol{v}_H = n\boldsymbol{c}$, where $n \geq 1$, so giving, for the present Hubble length:

$$L_H = nct_o/\sqrt{2}\,(\text{cm}) \tag{11}$$

4) The Hubble velocity of Space expansion may be obtained, as a function of time, from the above mentioned Equation (1) if $\boldsymbol{v}_H$ is substituted by $\dot{\boldsymbol{R}}$, as $\dot{\boldsymbol{R}} = \boldsymbol{\Gamma}_H \cdot t$ (1a) ($\text{cm}\cdot\text{s}^{-1}$).

4. The *c* Factor

It has been assumed [14] that the space expansion velocity could overpass $\boldsymbol{c}$ since the space is not in an inertial frame; rather, it contains all of reference frames. So, the concept of co-moving coordinates would be better applied to Space. From the above results, the t_c time has occurred at one fourth of the Universe age, near after the starting of the Physical Universe formation. That means that the light velocity into the Physical Universe has been evidently constant during the last three fourths of the Space age. There is a possible explanation for that: the expansion velocity of Space is not limited to $\boldsymbol{c}$ though matter velocity, into the Physical Universe, is really limited due to the space-time curvature originated by the mass density. However, to obtain a probable image of these subjects, there is no other way but to apply the available means [15] [16]. So it is possible to assume that, in the co-moving coordinates of Space, the expansion would be referred to the Big Bang origin ($t = r = m = 0$).

In order to determine the light velocity c as a function of the gravitational potential Φ, at a given distance, Einstein [17] did propose the following equation:

$$\boldsymbol{c}(\Phi) = \boldsymbol{c}_o\left(1+\frac{\Phi}{c_o^2}\right) \tag{12}$$

$\boldsymbol{c}_o$ is the present, known velocity. However, when applied to a higher potential corresponding to a smaller radius of the Physical Universe, the increase in $\boldsymbol{c}$ results almost negligible; it would confirm the constancy of $\boldsymbol{c}$ since the tiny obtained difference obeys, rather, to the imprecision of data applied in this paper.

Therefore, there are two realities: into the Physical Universe $\boldsymbol{c}$ is a true constant; in the Space, its expansion velocity pulls all matter (including photons) at higher velocities $\boldsymbol{v} > \boldsymbol{c}$. A possible reason for the $\mathbf{c}$ constancy could be the conservation of the angular momentum in the Physical Universe rotation that, at the t_c time, would have had a tangential velocity $\boldsymbol{c}$ which has been necessarily decreasing till the present time. So, if the total angular momentum of the Physical Universe ($\mathcal{L}$) is calculated at the t_c time, as:

$$\mathcal{L}_c = M_U \cdot r_c^2 \omega \tag{13}$$

by substitution of the relation $\boldsymbol{v}_c = \omega \cdot \boldsymbol{r}_c$ it gives:

$$\mathcal{L}_c = M_U \cdot \boldsymbol{r}_c \times \boldsymbol{v}_c \tag{14}$$

So, the constant value of the Physical Universe angular momentum at the t_c time it results: $\mathcal{L}$ = 4.8 × 10^{94} (erg·s); then, the to date tangential velocity of the Physical Universe Limit (R_{pu}), *i.e.* $z = 11$, would be $\boldsymbol{v}_{pu}^{t}$ = 2 × 10^{10} (cm/s).

It has been mentioned that a low gravitational potential has a very few influence in the case of an equilibrium temperature. However, in the non equilibrium temperature and a higher potential case, it seems that such influence would remain small.

5. The γ Factor

The Space expansion velocity (v_s) has really surpassed the $\boldsymbol{c}$ value after the t_c time and S. Hawkings [18] did mention the possibility of an imaginary time. Without a known theory about the empty space kinematics and too far from any gravitational interaction, the Special Theory of Relativity can be applied to obtain some expressions for a frozen and imaginary time, by means of the proper time (τ) equations:

$$\tau = t/\gamma \tag{15}$$

where:

$$\gamma = 1\bigg/\left(1-\frac{v^2}{c^2}\right)^{0.5} \tag{16}$$

i.e.

$$\tau = t\left(1 - v^2/c^2\right)^{0.5} \tag{17}$$

So, If $\boldsymbol{v} = \boldsymbol{c}$, $\gamma = \infty$, $\tau = 0$ (that means a time singularity). If $\boldsymbol{v} > \boldsymbol{c}$, τ has an imaginary value, τ. It may be determined the precise value of this imaginary number, by the γ factor: If $\boldsymbol{v} > \boldsymbol{c}$, it means that $\boldsymbol{v} = \beta\boldsymbol{c}$, being β a real number higher than 1, so $\beta^2 > 1$; if $w^2 = 1 - \beta^2$, $w^2 < 0$ and, from Equation (14),

$$\tau_i = t\left(w^2\right)^{0.5} = t \cdot wi \tag{18}$$

Therefore, when the velocity of some matter reaches $\boldsymbol{v} = \boldsymbol{c}$ and $\tau = 0$, that would mean a singularity (a time freezing at a physical Universe radius r_c), reached at one fourth of the present time. After that time ($t > t_c$), at an expansion velocity $\dot{R} > \boldsymbol{c}$, the proper time of space would become imaginary. Otherwise, as the physical Universe has maintained constant its maximum internal velocity $\boldsymbol{c}$, the proper time of photons and some leptons in the Physical Universe are really constant, *i.e.*, $\tau = 0$; that means that they remain into a singularity. That would be a strange situation into the physical Universe since the t_C time till today. Otherwise, the General Theory maintains the $\boldsymbol{c}$ limit for photons in the presence of mass; this one, in its turn, generates the curvature of space-time. The curvature decreases as the distance to a given point increases, similarly to the gravitational attraction does. At this point (r_c), the negative gravitational intensity $\boldsymbol{\Gamma}_G$ is 10^{-2} times lower than the positive expansion intensity $\boldsymbol{\Gamma}_H$.

6. The Thermo-Dynamical Expansion of the Universe

6.1. The Photonic Primeval Energy

The Big Bang has been assumed, from the G. Lemaitre concept, as a "primeval atom". Though it could not exist any type of atom at the Planck temperature, it really would mean a suddenly created energy accumulation, *i.e.* a photonic source; those photons applied to build the Physical Universe have generated a luminosity in the order of $\mathcal{L}_{pu} = 10^{59}$ (erg/s) [9]. Taking into account the assumed duration of the Big Bang of 10^{12} (s) ~ 3×10^4 (y), the total energy generated in the Physical Universe formation would have been about 10^{71} (erg). However, by including the original energy of CMB photons, a much higher value of the energy produced in the Big Bang could be obtained.

The original energy corresponding to the CMB can be obtained from the present density of such relic [19], *i.e.* 416 (photons/cm^3). By applying the cosmological principle, the total quantity of CMB photons in the Universe volume (7.4×10^{85} cm^3) should now be 3×10^{82} (photons). Since their average temperature is ~2.7 (K), it may be estimated an energy of 3.7×10^{-16} (erg/photon) which means a total present energy of 1×10^{67} (erg) for all CMB photons. From data of reference [20] it may be assumed that the z factor, for a frequency variation since the Big Bang to the CMB, would be ~10^{31}, (a similar z value may be deduced from the temperature variation) which means a total energy in the Big Bang period of 2×10^{98} (erg), equivalent to a luminosity $\mathcal{L}_B = 2 \times 10^{86}$ (erg/s). At the Big Bang end there would exist equal number of photons as they are today as CMB

(3×10^{82}), plus a small fraction devoted to generate the Physical Universe; so, the average energy of each one should have been $\sim E_f = 3 \times 10^{15}$ (erg/photon). However, the energy of just one photon whose maximum wavelength would equal the Planck length (as assumed here), it should have been about 1.5×10^{17} (erg/photon). That means that the average wavelength of original photons must rather be ~5 times longer than L_p, which implies a lowering factor of 5 in the original photons energy to give now 3×10^{16} (erg/photon) and a total energy at the Big Bang end of 2×10^{97} (erg). So, the luminosity at the Big Bang end would have been *Ļ* ~ 2×10^{85} (erg/s). This correction factor 5 in the wavelength of original photons doesn't match with the assumed Planck length. However, there are two possibilities for a coincidence: the first is that the "true" dimension of the Planck length would be 2.5 times bigger than the one here applied; the second possibility is that, accordingly to the quantum theory, a photonic wave is stable in a potential well if it displays an entire multiple of a half wavelength; that is a fact that permits to accept 5 different values of the ($\lambda/2$) parameter, so generating a mixture in the original energy spectrum of the Big Bang. Otherwise, the CMB fluctuations could have been provoked by dispersive interactions of original photons with the subsequent condensed matter.

Besides, the Higgs photon may have energies as high as ~150 GeV; that means ~0.25 (erg/photon). So, it would be feasible that every one of the original photons could generate the necessary number of Higgs photons to start the imminent mass condensation.

The total energy was, apparently, applied to only two purposes: one was the generation of the Physical Universe mass, whose value results $\sim 2 \times 10^{78}$ (erg) if the mass would have a value $\sim 2 \times 10^{57}$ g [21]. It means that a minimal fraction (10^{-19}) of the Big Bang energy was applied to generate the Physical Universe. Accordingly to reference [9] a similar fraction would have been applied to the Physical Universe expansion work. The rest (>99.9%) was left as the relic of the Big Bang. It would imply that the present electromagnetic spectrum, from radio to gamma and cosmic rays, as well as some leptons, would not come from the Big Bang but from astronomical objects such as the Sun, galaxies and quarks.

6.2. The Expansion Work of the Physical Universe

In this expansion, temperature, pressure and volume continuously vary, though entropy does not because there is not any intake or lost of heat in the Universe. The conservation equation for the internal energy of the Universe (E) is, from the 1st Law of Thermodynamics:

$$E = Q - W \left(\text{erg}\right) \tag{19}$$

Q would be the total heat content, and W the work performed by the matter expansion. Initially, at a time lower than Planck's one, the work expansion W may be assumed to be nil and therefore the internal energy U would be equal to the heat content ($E = Q$). After, in the non-isothermal, non isobaric and adiabatic (though irreversible) case of the Physical Universe, the expansive work will be made at expenses of a small fraction of the original internal energy: ($W = -\delta E$). So, E is the original internal energy of the Big Bang, which was above calculated as 2×10^{97} (erg); δE_M, corresponding to the mass generation, is 2×10^{78} (erg), as previously mentioned. By derivation of Equation (19) respect to time and substitution of the density variable, the conservation equation for adiabatic expansion it results:

$$\dot{\rho} + 3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) = 0 \tag{20}$$

where a and $\dot{a}$ refer to the scale factor and its time derivative.

However, additional considerations must be made: accordingly to reference [10] "it cannot be pressure forces in a homogeneous universe because such effects can only be generated by a pressure gradient… so, pressure does not provide a force that causes the Universe expansion. Rather, its contribution is entirely through the work done during the expansion". So, it has not sense to search for a pressure value in the Universe expansion process. Otherwise, the state equation for different types of matter has eventually been proposed as:

$$p = w\rho c^2 \tag{21}$$

For vacuum energy, $w = -1$, so assuming that an external negative pressure is the cause of Universe expansion. Such possibility has been discarded by both Λ and Hubble models.

There are three ways to determine the work developed in the Physical Universe expansion: to apply the Hubble potential, by means of a classical thermo-dynamical process, and to include a quantum criteria.

1) The Hubble potential acts on the masses present after the decoupling time; so, the to-date work value could be:

$$W = M_U \Gamma_H R_X \ (\text{erg}) \tag{22}$$

M_u is the mass of the physical Universe and R_x is the radial function corresponding to the physical Universe expansion period, $t_x = (t_o - t_{pu})$. Substitution of known values gives a total expansion work of the physical Universe $W_U \sim 10 \times 10^{76}$ (erg) that was delivered at the spherical shell R_x, *i.e.* during the evolution time of the physical Universe, which represents a luminosity of $L_, = 2 \times 10^{59}$ (erg/s), similar to that of reference [9] for the physical Universe. The negative external pressure would be 0 since the empty space density is 0 (different to that of a quantum vacuum). Anyway, the sum of the Physical Universe mass and its expansion work, remain in a fraction lower than 10^{-19} of the Big Bang energy.

2) The classic thermo-dynamical expansion work is given by the equation:

$$W = PV_u = nkT \ \left(\text{dyne}/\text{cm}^2\right) \tag{23}$$

where, n has been calculated as 2.7×10^{82} (photons), $k = 1.38 \times 10^{-16}$ (erg/K·photon); if $T_u \sim 2 \times 10^{10}$ (K), when the lepton era ends, the product gives $PV = 7.5 \times 10^{76}$ (erg).

3) The quantum method [22] applies the Wien factor b = 7.6 × 10^{-15} (erg/K·cm^3), obtained from statistical physics, to get the thermal pressure; however, since it was derived for an isothermal process, it may not be here applied to an adiabatic expansion. Anyway, the quantum method includes a correction factor for the classical Equation (23) that was here employed, so arriving to a result W_U = 6.7 × 10^{76} (erg). The luminosity, in this case, is similar to that corresponding to the evolution of the physical Universe: Ļ ~ 1.3 × 10^{59} erg/s.

6.3. A Big Bang Design

Before making additional calculation in the photonic gas, it must be recalled the classical Bose-Einstein statistics for 0 spin particles in a non-isothermal process, *i.e.*

$$f(E) = 1/\left(Ae^{E/kT} - 1\right) \tag{24}$$

that is applied to obtain the distribution of photons as a function of temperature [23]; it may not be used here because it represents an equilibrium state at a given temperature T, when photons energies vary and most photons accumulate at the lowest energies; it is an ideal though opposed situation to that of the Universe where the photons accumulation happened at the highest temperature. Therefore, it is possible to assume that the only reducing factors of the initial photons energy will be a gravitational (delaying) red-shift and, after the decoupling time, the normal z lowering process driving to the CMB.

In order to prove the initial conditions for the validity of a physical law, it would be important to calculate the expansion velocity of the Big Bang in comparison to that of the Hubble acceleration. So, the general Equation (23) is assumed to represent the adiabatic expansion work of the photonic nucleus; then:

$$\left(\dot{R}\right)^2 = 2nkT/M_{pu} \tag{25}$$

So, $\dot{R}$ = 5 × 10^6 (cm·s^{-1}) would be the final velocity of the Big Bang period, while the velocity obtained from the Hubble acceleration results: =2.7 × 10^5 (cm·s^{-1}). This difference must have been much higher at shorter times, which means that the Hubble acceleration law was not significant before the end of the Big Bang period (3 × 10^4 y). Therefore, this period could be assumed as the one corresponding to inflation. Taking the Planck length L_p = 1.6 × 10^{-33} (cm) as the diameter of an spherical Big Bang, its volume would be V_p = 2.15 × 10^{-99} (cm^3).

In **Table 1**, they are shown the values of expansion velocity $\dot{R}_H$ (cm/s) and scale factor R_H (cm) of the Space, corresponding to times (s): Planck (t_p); 10^{-10}; 1; Big Bang end, t_{be}; decoupling time t_d; Physical Universe formation time t_{pu}; $\boldsymbol{c}$ time, t_c; and present times: t_{11} for the Physical Universe and t_o for Space. Four values for luminosity Ļ (erg/s) are also shown for times t_{be}, t_d and t_o times, as well as to the period (t_c − t_{be}). The present scale factor of the Physical Universe is assumed as the observable Universe: the distance to the most distant object (GN-Z-11), R_{11} ~ 2 × 10^{27} (cm).

Table 1. Expansion velocity (v), and scale factor (R) of the Space, according to the Planck parameters (first line) and to the Hubble acceleration (next 8 lines). The luminosity $Ļ$ (erg/s) of Space is shown in the fourth column for: the Big Bang end (t_{be}), the decoupling time (t_d), the (t_c) time and the present time (t_o).

t (s)	$\dot{R}$ (cm/s)	R (cm)	$Ļ$ (erg/s)
$t_p = 5.4 \times 10^{-44}$	c	$L_p = 1.6 \times 10^{-33}$	
$t = 10^{-10}$	2.7×10^{-17}	$R_{-10} = 1.4 \times 10^{27}$	
$t = 1$	2.7×10^{-7}	$R_1 = 1.4 \times 10^{-7}$	
$t_{be} = 10^{12}$	2.7×10^{5}	$R_{be} = 1.5 \times 10^{17}$	$Ļ_{be} = 2 \times 10^{85}$
$t_d = 1.5 \times 10^{13}$	3.7×10^{6}	$R_d = 4.6 \times 10^{19}$	$Ļ_d = 6 \times 10^{83}$
$t_{pu} = 3.1 \times 10^{16}$	$0.3c$	$R_{pu} = 3.0 \times 10^{24}$	
$t_c = 1.1 \times 10^{17}$	c	$R_c = 1.6 \times 10^{27}$	$Ļ(t_c - t_{be}) = 4.5 \times 10^{80}$
$t_{I1} = t_o - t_{pu}$	$3c$	$R_{11} = 2.0 \times 10^{27}$	
$t_o = 4.4 \times 10^{17}$	$4c$	$R_o = 2.6 \times 10^{28}$	$Ļ_o = 5 \times 10^{79}$

7. The Photonic Equation of Continuity

The operator of the convective derivative has been usually applied to density and time variables such as velocity, momentum and kinetic energy. The same operator may also be applied to these parameters when they are not expressed as function of masses but of frequencies, such in the photons case. Applying the convective operator to the photon energy, $E = h\nu$, if $E(r, t)$, it gives:

$$Dh\nu/dt = \partial h\nu/\partial t + \boldsymbol{c}\partial h\nu/\partial r \tag{26}$$

Substitution of $\boldsymbol{c}$ in ∂r drives to the equation of continuity:

$$\frac{DE}{dt} = 2h\left(\frac{\partial \nu}{\partial t}\right) \tag{27}$$

This rather unexpected result may be applied to an example with the above mentioned parameters if $h\partial\nu$ would be assumed as $\Delta E = (E_c - E_{be})$ and ∂t as $\Delta t = (t_c - t_{be})$ to give:

$$\frac{DE}{dt} = 2\left[\frac{\Delta E}{\Delta t}\right](\text{erg/s}) \tag{28}$$

where, the sub-fix $_{\text{be}}$ refers to the Big Bang end time (t_{be}) and $_{\text{c}}$ to the t_c time. Equation (22) gives a result $DE/dt = 4.5\times10^{80}$ (erg/s); it would be the rate of outgoing energy, since the Big Bang end time till the t_c time. This is the luminosity (L_c) of the Universe corresponding to the period ($t_c - t_{be}$). The Jacobian for this energy in adiabatic expansion may be expressed as:

$$J(E,S) = \frac{J(E,S)}{J(t,S)} = \left(\frac{\partial E}{\partial t}\right)_s \tag{29}$$

Substitution gives a similar result:

$$(DE/dt)_s = 4.3\times10^{80}\ (\text{erg}) \tag{30}$$

8. Conclusions

1) It has been assumed that, at the Big Bang, there were created four realities: time, space, matter and physical laws. Several references [24], [25], [26] ask about the time creation of the physical laws. If the above mentioned constancy of the Space acceleration $\mathbf{\Gamma}_H$ is correct, it would represent the first physical law created at the Big Bang, respect to Space, together with the gravitational law $\mathbf{\Gamma}_G$, respect to matter; however, $\mathbf{\Gamma}_H$ would appear evident only after the Big Bang end, as well as $\boldsymbol{c}$ would be freely manifest till the t_c time. Additional parameters, such as Luminosity, are summarized in **Table 1**. It is also assumed that the Universe is the total accelerating Space that contains the Physical Universe and an external vacuum Space.

2) The Big Bang was not an instantaneous event; it remained for 10^{12} (s) or 3×10^4 (y) [9]. Given both the length and the expansion velocity at this stage, it may be assumed to correspond to the Inflation period.

3) The $\boldsymbol{c}$ time (t_c) was defined as that when Space and the Physical Universe reached together the **c** value of expansion velocity; at such time, both Space and Physical Universe had a radius $\boldsymbol{r}_c$. In spite the gravitational law had continuously been opposed to the Physical Universe expansion, the Space expansion has independently occurred due to the constancy of the Hubble acceleration expressed by $\mathbf{\Gamma}_H$. However, inside the Physical Universe, matter velocity has kept a maximum velocity $\boldsymbol{c}$ as a universal constant, a consequence of the space-time curvature. So, the furthest tangential velocity of the Physical Universe, at the t_c time, could also have been $\boldsymbol{c}$, an assumed data that had permitted to calculate the Physical Universe angular momentum as $\mathcal{L} = 4.8 \times 10^{94}$ (erg·s).

4) According to Special Relativity, at **c** velocity the proper time becomes nil, *i.e.* $\tau = 0$, which implies a time freezing, as well that matter felt in a singularity; it would be difficult to define the meaning and length of such singularity. Also, it has been mentioned [18] the theoretical possibility of an imaginary time (τ_i) if the limits of the Physical Universe overpasses the $\boldsymbol{c}$ velocity. It has been shown in a simplified Euclidic Space[3] [26].

5) The to-date called Hubble Length, $L_H = \boldsymbol{c}/H$, had been defined as the distance where the space expansion rate $\dot{R}$ is just $\boldsymbol{c}$; so, farther L_H, $\dot{R} > \boldsymbol{c}$. However, H is usually assumed as constant though it should rather be determined as H_o, *i.e.* a function of the present scale factor R_o, obtained from the constant value of $\mathbf{\Gamma}_H$. Besides, the H_o value deduced in reference [1] has been modified in this paper to $H_o = 3.2 \times 10^{-18}$ (s^{-1}); the value of $\mathbf{\Gamma}_H$ is 2.65×10^{-7} ($\text{cm}\cdot\text{s}^{-2}$).

6) The parameter Ω has no sense at the expanding limit of the Physical Universe since the density of the external space is 0. As well, the q parameter must be >0 in an expanding Physical Universe, because every one of its factors is >0. Respect to the Hubble parameter, it was here deduced the equation $H(t) = \sqrt{2}/t$ (8) for an expansive Universe.

---

[3] The Euclidic name has been proposed by S. Hawkings for a Euclidean Space that includes an imaginary time coordinate [18].

7) The expansive work calculation was intended by 3 ways: the Hubble potential, the thermo-dynamical process and a brief quantum concept. They gave $W \sim 7.4 \times 10^{76}$ (erg).

8) The Big Bang was a source of photons, staying for 3×10^4 years to liberate a total energy of 2×10^{97} (ergs) at a constant true velocity $\boldsymbol{c}$ that seemed impossible into the Big Bang tangle. A minimum part of this energy ($\sim 10^{-21}$) was applied to build and expand the Physical Universe; the rest of photons decayed necessarily till the present known CMB energy, without any foreseeable task.

9) The convective operator was applied to photons energy, as a function of frequency, to get an Equation (27) derived from the assumption that energy is a function of time and distance. As well, the Jacobian obtained for the adiabatic expansion gave a similar result for Luminosity.

10) The above calculated initial energy of photons is high enough to generate the necessary number of Higgs photons in order to produce mass condensation. The above mentioned totality of initial photons would not be mono-energetic; their variation is tied, between several factors, to the Heisenberg indetermination of the Planck length, which results ~5 times the proper length. Fluctuations of CMB photons could be originated by dispersive interactions with the Physical Universe matter.

11) Some conclusions require the experimental determination of the true Space acceleration expansion, here deduced as $\boldsymbol{\Gamma}_H = 2.65 \times 10^{-7}$ (cm·s^{-2}).

Acknowledgements

The author thanks to M. A. Zúñiga for his kind collaboration in this work.

Conflicts of Interest

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

References

[1] Lartigue, J.G. (2016) *Journal of Modern Physics*, **7**, 1607-1615. https://doi.org/10.4236/jmp.2016.712145

[2] Mather, J.C. (2006) From the Big Bang to the Nobel Prize. Lecture.

[3] Ford, I. (2013) Statistical Physics. J. Wiley, New York, 187-199. https://doi.org/10.1002/9781118597507

[4] Pacavic, A. (1963) Thérmodynamique Statistique. Masson et Cie, Paris, 238.

[5] Chen, J. (2010) The Relativistic Dynamics of a Photon Gas. p. 6.

[6] Bergstrom, L. and Goobar, A. (2004) Cosmology and Particle Astrophysics. 2nd Edition, Springer, Berlin, 60.

[7] Lartigue, J.G. (2017) *Journal of Modern Physics*, **8**, 6, Eq. (4.9).

[8] Salvo, R., *et al.* (2017) *Science*, **358**, 765-768. https://doi.org/10.1126/science.aan4054

[9] Hogan, C.J. (1999) Energy Flow in the Universe. <arXiv:astro-ph/9912110v|>.

[10] Lidsey, J.E. (2014) Cosmology Course ASTM108/PHY7010U, eq. (4.2), Queen Mary School, London.

[11] Lartigue, J.G. (2016) *Journal of Modern Physics*, **7**, 1607-1615, eq. (3.3). https://doi.org/10.4236/jmp.2016.712145

[12] Sartory, L. (1996) Understanding Relativity. University of California, Los Angeles, 313.

[13] Lidsey, J. op. cit. eq. (2.28).

[14] Sartory, L.(1996), op. cit. , p. 325, Figure 9-12.

[15] Gleiser, M. (2013) The Origin of the Universe. https://www.npr.org/section/13.7

[16] Tolman, R. (1958) Relativity, Thermodynamics and Cosmology. Oxford Press, Oxford, 487.

[17] Einstein, A., *et al.* (1923) The Principle of Relativity. Dover Pub., 106.

[18] Hawkings, S. (1988) Historia del tiempo. Editorial Crítica, México, 179.

[19] Leff, H.S. (2002) *American Journal of Physics*, **70**, 792-797. https://doi.org/10.1119/1.1479743

[20] Sartory, L. (1996), op. cit. p. 342, Table 9.1.

[21] Johnson, J.R. (2013) Comprehending the Cosmos. 2nd Edition, Create Space Independent Publication, 204, 264.

[22] Leff, H. S. (2002) op. cit. p. 793.

[23] Krane, K. (1983) Modern Physics. Wiley, New York.

[24] Davies, P. (1995) L'Esprit de Dieu. Éditions du Rocher, Paris, 83.

[25] Hawkings, S. (op. cit.) p. 164.

[26] Lartigue, J.G. (2018) The Universe Origin and Destiny. *Journal of Modern Physics*, **9**, 1823, Figure 3.

Journal of Modern Physics, 2018, 9, 2457-2472
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Ab-Initio Computations of Electronic, Transport, and Related Properties of Chromium Disilicide ($CrSi_2$)

Shaibu Onuche Mathias, Yuriy Malozovsky, Lashounda Franklin, Diola Bagayoko

Department of Mathematics and Physics, Southern University and A & M College, Baton Rouge, LA, USA
Email: cecemat4all@yahoo.com

How to cite this paper: Mathias, S.O., Malozovsky, Y., Franklin, L. and Bagayoko, D. (2018) *Ab-Initio* Computations of Electronic, Transport, and Related Properties of Chromium Disilicide ($CrSi_2$). *Journal of Modern Physics*, **9**, 2457-2472.
https://doi.org/10.4236/jmp.2018.914158

Received: November 17, 2018
Accepted: December 17, 2018
Published: December 20, 2018

Abstract

We report results from *ab-initio*, self-consistent density functional theory (DFT) calculations of electronic, transport, and related properties of chromium disilicide ($CrSi_2$) in the *hexagonal C*40 crystal structure. Our computations utilized the Ceperley and Alder local density approximation (LDA) potential and the linear combination of atomic orbitals (LCAO) formalism. As required by the second DFT theorem, our calculations minimized the occupied energies, far beyond the minimization obtained with self-consistency iterations with a single basis set. Our calculated, indirect band gap is 0.313 eV, at room temperature (using experimental lattice constants of a = 4.4276 Å and c = 6.368 Å). We discuss the energy bands, total and partial densities of states, and electron and hole effective masses. This work was funded in part by the US Department of Energy, National Nuclear Security Administration (NNSA) (Award No. DE-NA0003679), the National Science Foundation (NSF) (Award No. HRD-1503226), LaSPACE, and LONI-SUBR.

Keywords

Band Gap, BZW-EF Method, Density Functional Theory, Band Structure, $CrSi_2$

1. Introduction and Motivation

Chromium disilicide, $CrSi_2$, belongs to a list of semiconducting metal-silicides. It has a C40 hexagonal crystal structure, with a space group of $P6_222$ [1] [2] [3] [4], as depicted in **Figure 1(b)**. It is a highly degenerate p-type semiconductor with a narrow-forbidden band gap [5] [6]. $CrSi_2$ exists in several compositions [7] [8], ranging from 65.7% to 67.7% silicon [9]. The compound has three (3) formula

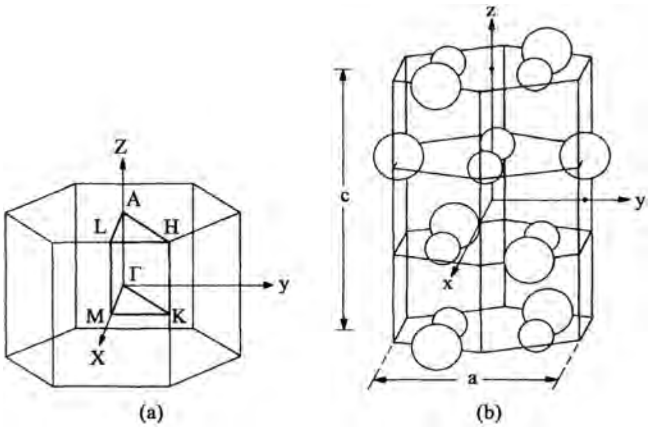


Figure 1. (a) Brillouin Zone for Hexagonal $CrSi_2$ and (b) Primitive unit cell of $CrSi_2$. Large spheres represent Cr atom positions while small spheres represent Si atom positions.

units per hexagonal unit cell [9]. It belongs to a group of semiconducting metal-silicides which have gained enormous attention in recent years, due to its properties and several areas of important applications. It has been the most studied representative of the metal-silicides since its initial characterization as a 0.35-eV bandgap semiconductor in the mid-1960's [10]. Due to the semiconducting nature and thermal stability of $CrSi_2$, it has special applications in optoelectronic devices, infrared detectors within silicon-based microelectronics components [3] [9] [11] [12]. $CrSi_2$, as a high-temperature compound, has been epitaxially grown on Si (111) substrate [1] [4] [13]. The preceding property of $CrSi_2$ makes it a potential material in the production of thermoelectric generators as well as for photovoltaic applications, in the middle of the infrared region [14]. As a narrow band gap semiconductor, $CrSi_2$ is a very good candidate in micro- and nano-electronics, respectively, and for photo-thermo converters and sensors [4] [15]. $CrSi_2$ belongs to a group of refractory silicides with a melting point at 1763 K, which makes it a potential candidate for high-temperature applications. $CrSi_2$ films are widely used in the area of new semiconductor device manufacturing due to their excellent electronic properties, high thermal stability, smooth surface and remarkable compatibility with the traditional silicon technique [12]. Krivosheeva *et al.* [16] reported that one of the most interesting and well investigated compounds is chromium disilicide which has the smallest lattice mismatch, as compared to other transition metal silicides [17] [18], with mono-crystalline silicon. $CrSi_2$ has a high electrical conductivity and a strong oxidation resistance which make it more attractive in microelectronics [12]. $CrSi_2$ is a potential candidate for optoelectronic devices, photo-voltaic cells, and thermoelectric conversion elements operating at elevated temperatures [1] [17] [19] [20] [21] [22].

Some experimental data have been reported for hexagonal $CrSi_2$. However, a consensus has not been reached, as far as its band gap is concerned; one reason for this situation stems from the lack of measured band gap values for bulk

$CrSi_2$. Bost *et al.* [9], in optical studies on well characterized $CrSi_2$ polycrystalline thin films, in 1988, obtained experimental results that provide evidence for the semiconducting nature of $CrSi_2$. Results from their measurements showed that $CrSi_2$ exhibits an indirect band gap of 0.35 eV [9]. Additionally, in a study of optical properties of $CrSi_2$, Henrion *et al.* [23] reported a band gap of 0.50 eV for $CrSi_2$ polycrystalline thin films, in 1992. Experimental studies of $CrSi_2$ films synthesized by high current Cr ion implantation resulted in band gap values of 0.7 eV and 0.84 eV [24] for $CrSi_2$ layers under different experimental conditions. Energy band gaps of 0.30 eV [25] to 0.35 eV [26] were obtained for $CrSi_2$ from Hall-effect measurements. Nishida [27] measured a band gap of 0.32 eV for $CrSi_2$ single crystals grown by using the floating zone melting technique. This author did not state whether the measured band gap was direct or indirect. Results from ellipsometry [28] suggested an indirect band gap of Eg ≤ 0.36 eV for $CrSi_2$. All of the experimental band gaps reported so far for $CrSi_2$ are indirect except for the work of Nishida and of Galkin *et al.* [29]. While the former did not specify the nature of the gap, the latter found a direct band gap of 0.37 eV, for $CrSi_2$ epitaxial films. Clearly, results provided from past experimental works are not in total agreement. However, a general consensus points to a band gap in the range of 0.27 - 0.8 eV for various films of $CrSi_2$. **Table 1** shows experimental band gap values reported for $CrSi_2$.

Table 1. Results from Experimental Measurements of the Band Gap of Hexagonal $CrSi_2$. Except for the one indicated to be direct, all band gaps below are indirect. These band gaps are for films of various thicknesses, except the 0.32 eV band gap value estimated from the temperature dependence of resistivity for a single crystal.

Growth Or Measurement Method	E_g (eV)
Laser-assisted Synthesis of semiconductor chromium disilicide films	0.2^{a}
Polycrystalline samples grown by amorphous thin films of Cr and Si in double electron-gun evaporation system.	0.27 ± 0.01^{b}
Hall Effect measurements of Si-doped and Mn-modified $CrSi_2$ crystal	$0.30 - 0.35^{c}$
Single crystals of $CrSi_2$ grown using the floating zone melting technique. Energy gap estimated from the temperature dependence of resistivity.	0.32^{d}
Synchrotron Radiation Photoemission measurement of epitaxial $CrSi_2$ films prepared on Si (111) substrate at room temperature and 20K	0.32^{e}
$CrSi_2$ films prepared by molecular beam epitaxy on $CrSi_2$ templates grown on Si (111) Substrate	0.34^{f}
Polycrystalline thin films of $CrSi_2$ grown on silicon substrates (Samples annealed at 1100˚C)	0.35^{g}
Ellipsometry of polycrystalline thin films of $CrSi_2$	$\leq 0.36^{h}$
Optical absorption measurement of $CrSi_2$ thin films	$0.35 - 0.5^{i}$
Transmittance and Reflectance Spectroscopy Study of A-type Epitaxial films 100nm thick grown by the Template method	0.37^{j} direct
Optical Spectra measurement of $CrSi_2$ polycrystalline thin films	0.50^{k}
Synthesis of $CrSi_2$ films by high current Cr ion implantation	0.7 and 0.8^{l}

[a]Ref. [15], [b]Ref. [28], [c]Ref. [25] [26], [d]Ref. [27], [e]Ref. [30], [f]Ref. [31], [g]Ref. [9], [h]Ref. [28], [i]Ref. [23] [32], [j]Ref. [29], [k] Ref. [10], [l]Ref. [24].

Several theoretical calculations have been reported for the electronic structure of $CrSi_2$. While some of the calculations [31] have argued that $CrSi_2$ is semi-metallic in nature, others have predicted semiconductor properties for this material. Dasgupta *et al.* [4] obtained an indirect band gap of 0.35 eV, using the augmented spherical wave (ASW) method [33] [34] and the generalized gradient approximation (GGA) potential parameterized by Perdew *et al.* [35]. However, another calculation [20] performed with a similar method led to indirect and direct band gaps of 0.21 eV and 0.39 eV, respectively. Bellani *et al.* [28] reported a theoretical indirect band gap value of 0.38 eV using the linear-muffin-tin-orbital (LMTO) method, within the local density approximation (LDA). Two (2) calculations [19] [36] using the same method, within the local density approximation (LDA), obtained indirect band gaps of 0.29 eV and 0.25 eV, respectively. Another calculation [37], utilizing the LMTO method within the atomic spheres approximation (ASW), obtained a gap of 0.38 eV. L. F. Mattheiss [11] [38] reported an indirect band gap of 0.30 eV for bulk $CrSi_2$, using the linear augmented plane wave method (LAPW) and a local density approximation (LDA) potential. Mattheiss [39] used a scalar-relativistic version of the linear augmented-plane-wave (LAPW) method and obtained an indirect band gap of 0.30 eV. In another DFT calculation [16], with the full-potential linearized-augmented-plane-wave (FP-LAPW) led to an indirect band gap of 0.30 eV. A DFT approach, similar to the preceding, was applied in another calculation to obtain a band gap of 0.30 eV [10]. Zhou ShiYun *et al.* [12] obtained a gap of 0.353 eV in their study of optical properties of $CrSi_2$; they employed the plane-wave pseudo-potential method. Finally, recent DFT calculations performed in 2013 by Bhamu *et al.* [40] produced an indirect band gap of 0.28 eV for $CrSi_2$. The above calculation methods, potentials, and results are listed in **Table 2**.

Many of the results obtained from both experimental and theoretical calculations of $CrSi_2$ have been extensively reviewed in the preceding section. It is clear, however, from the contents of **Table 1** and **Table 2** that these results do not totally agree. While the disagreement can be seen among theoretical results, on the one hand, and between experimental results, on the other hand, there exists also a disagreement between experimental and theoretical results. This disagreement between theoretical results can be partly attributed to differences in computational methods. These disagreements strongly suggest that the correct band gap of bulk $CrSi_2$ is yet to be established unambiguously. This situation is a key motivation for our work. Also, the many current and potential applications of $CrSi_2$, as discussed at the beginning of this section, also motivated this work. These two motivations are supported by the fact that our method, to be discussed below, has led to the correct band gaps of well over 30 semiconductors. This method correctly predicted the band gap and related properties for more than three (3) semiconductors. Our aim, therefore, is to obtain accurately, through our BZW-EF, *ab-initio* self-consistent calculations, the true band gap as well as other electronic, transport and related properties of $CrSi_2$. Our BZW-EF *ab-initio*, self-consistent method has been successfully applied in several calculations

Table 2. Results from previous theoretical calculations of the band gap of hexagonal $CrSi_2$. except for the one indicated to be direct, all the gaps in the table are indirect.

Computational method	**Potentials**	**E_g (eV)**
Augmented-Spherical-Wave (ASW)	LDA	0.21[a]
Linear-Muffin-Tin-Orbital (LMTO)	LDA	0.25[b]
Linear Combination of Atomic Orbitals (LCAO)	LDA	0.28[c]
Semi-relativistic Linear Muffin-Tin-Orbital	LDA	0.29[d]
Linear Augmented-Plane-Wave (LAPW)	LDA	0.30[e]
Scalar Relativistic Linear Augmented-Plane-Wave	LDA	0.30[f]
Plane-Wave Pseudopotential Theory	LDA	0.353[g]
Full-Potential-Linearized-Augmented-Plane Wave (FL-APW)	LDA	0.35[h]
Semi-linear theory of relativity of the linear-muffin-tin-orbital (LMTO)	LDA	0.38[i]
Linear Muffin Tin Orbital (LMTO)	LDA	0.38[j]
Linear Augmented Plane Waves (LAPW)	GGA	0.30 (direct)[k]
Full-Potential-Linearized-Augmented-Plane Wave (FP-LAPW)	GGA	0.30[l]
Augmented Spherical Wave (ASW)	GGA	0.35[m]

[a]Ref. [20], [b]Ref. [19], [c]Ref. [40], [d]Ref. [36], [e]Ref. [11] [38], [f]Ref. [39], [g]Ref. [12], [h]Ref. [37], [i]Ref. [28], [j]Ref. [28], [k]Ref. [10], [l]Ref. [16], [m]Ref. [4].

[41]-[52] in the past and has proven to produce accurate properties of semiconductors. Therefore, this work is expected to follow in the same light.

2. Our Distinctive Method and Computational Details

Our computational method has been extensively discussed in previous publications [41]-[49], [53] [54] [55]. Two components of this method are commonly utilized in most calculations, *i.e.*, the choice of a density functional potential (LDA or GGA) and the linear combination of atomic orbitals (LCAO). Our software package actually employs the linear combination of Gaussian orbitals (LCGO). We selected the LDA potential of Ceperley and Alder, as parameterized by Vosko *et al.* [56] [57].

The distinctive feature of our method consists of our utilization of successive, self-consistent calculations, with augmented basis sets, in order to minimize the energy content of the Hamiltonian. This process ultimately leads to the absolute minima of the occupied energies (*i.e.*, the ground state), as required by the second theorem of density functional theory. This feature in our calculations is known as the Bagayoko, Zhao, and Williams (BZW) method [41] [58]-[63], as enhanced by Ekuma and Franklin (BZW-EF) [47] [48] [49] [54]. Unlike the BZW method, where orbitals representing unoccupied states are added in the order of increasing energies (in atomic or ionic species), the enhanced version (BZW-EF) adds, for a given principal quantum number, p, d and f orbitals, when applicable, before adding the corresponding s orbital. An orbital is applicable if it is occupied in any of the atomic species in the system. The BZW-EF method reflects the realization [46] that polarization orbitals, for valence electrons, have primacy over the spherically symmetric s orbital [46] [47] [48] [49] [53] [55]. We describe below the actual implementation of the method using the program package developed at the Ames Laboratory of the US Department of Energy (DOE), Ames, Iowa [64] [65].

Our calculations for $CrSi_2$ started with a small basis set that was not smaller than the minimum basis set. This first self-consistent calculation was followed by Calculation II whose basis set was that of Calculation I as augmented with one orbital representing an excited state. Every augmentation of the basis set increases the dimension of the Hamiltonian by 2, 6, 10, or 14, depending on the s, p, d, or f character of the added orbital, respectively. We compared the self-consistent eigenvalues of the two calculations, graphically and numerically. Some occupied energies from Calculation II were lower than corresponding ones from Calculation I, as expected. After augmenting the basis set of Calculation II, Calculation III was performed self-consistently. The comparison of the occupied energies of Calculations II and III showed that some occupied energies of Calculation III were lower than corresponding ones from Calculation II. This process continued until three (3) consecutive calculations led to the same occupied energies, within our computational uncertainty of 5 meV, indicating that the ground state was reached. The first of the three (3) consecutive calculations was selected as the one providing the DFT description of the material; the basis set of this calculation is referred to as the *optimal basis set* [49]. As shown in the Section on results, this calculation was Calculation IV that produced the same occupied energies as V and VI. *The selection of the optimal basis set in the BZW-EF method is based on the crucial fact that the charge density from this calculation is the same one obtained in the calculations following it*. Hence, the Hamiltonian for this calculation, in light of the first theorem of DFT, is the same as those calculations following it, even though the Hamiltonian matrices will be different, given their different dimensions. Bagayoko [53] explained the reason the calculation with the optimal basis set is the one providing the DFT description of the material. Self-consistent iterations, up to the calculation producing the optimal basis set, yield eigenvalues that are due to interactions in the Hamiltonian. Calculations with basis sets larger than the optimal one and that contain the optimal one do not change the Hamiltonians or the occupied energies from their respective values obtained with the optimal basis set. However, these calculations can produce unoccupied energies that are lower than their corresponding values obtained with the optimal basis set. Given that the Hamiltonians of these calculations are the same as that obtained with the optimal basis set, the unoccupied energies lowered below their values obtained with the optimal basis no longer belong to the spectrum of the Hamiltonian, a unique functional of the charge density [53].

Computational details for this work follow. Chromium disilicide ($CrSi_2$) has a hexagonal C40 structure. It is in the space group of $P6_222$ $\left(D_6^4\right)$ [1] [2] [3] [4].

Its primitive cell contains a total of three (3) CrSi_2 formula units with individual atoms arranged as shown in **Figure 1(b)**. The space group is non-symmorphic, containing non-primitive translations ($\tau = \frac{c}{3}$ and $\frac{2c}{3}$) which interchange individual CrSi_2 layers [11] as in **Figure 1(b)**. Each Cr and Si atom in each of hexagonal layers of CrSi_2 has six (6) nearest neighbors at d = 2.557 Å. Each Cr and Si atom also has four (4) interplanar neighbors which are tetrahedrally coordinated. The hexagonal Bravais lattice for the primitive cell of CrSi_2 is generated from the primitive vectors: $\boldsymbol{t}_1$, $\boldsymbol{t}_2$ and $\boldsymbol{t}_3$, each described in Equation (1) as

$$\boldsymbol{t}_1 = (a/2)\left(\sqrt{3}\hat{i} - \hat{j}\right),\ \boldsymbol{t}_2 = a\hat{j},\ \boldsymbol{t}_3 = c\hat{k} \tag{1}$$

where a and c are the lattice constants. The internal atom position coordinates (ξ, ζ, η) for the primitive unit cell of CrSi_2 are in the units of the primitive vectors in Equation (1). These position coordinates of Cr and Si, within the hexagonal C40 primitive unit cell of CrSi_2, are given in **Table 3**, where x is the Si-atom position parameter. The position parameter of the Si-atom does not have an exact value. However, a value of $x = 1/6$ [64], corresponding to an ideal geometry [11], is normally used. In the ideal geometry, each Cr and Si atom has six nearest neighbors (d = 2.55 Å) [11], within each hexagonal CrSi_2 layer.

The standard hexagonal Brillouin zone for CrSi_2, as shown in **Figure 1(a)**, was generated from the reciprocal-lattice vectors that correspond to Equation (1). These reciprocal-lattice vectors are described by Equation (2) as given below.

Table 3. Position coordinates (ξ, ζ, η) of Cr and Si atom within the primitive unit cell of hexagonal C40 CrSi_2 in units of primitive vectors.

Atom	Site	ξ	η	ζ
Cr	3d	$\frac{1}{2}$	0	$\frac{1}{2}$
		0	$\frac{1}{2}$	$\frac{1}{6}$
		$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{6}$
Si	6j	x	$2x$	$\frac{1}{2}$
		$-x$	$-2x$	$\frac{1}{2}$
		$2x$	x	$\frac{1}{6}$
		$-2x$	$-x$	$\frac{1}{6}$
		x	$-x$	$-\frac{1}{6}$
		$-x$	x	$-\frac{1}{6}$

$$\boldsymbol{b}_1 = \left(4\pi/\sqrt{3}a\right)\hat{i},\ \boldsymbol{b}_2 = \left(2\pi/\sqrt{3}a\right)\left(\hat{i}+\sqrt{3}\hat{j}\right),\ \boldsymbol{b}_3 = \left(2\pi/c\right)\hat{k}, \qquad (2)$$

where a and c are the lattice constants.

Our non-relativistic, self-consistent calculations were performed using room temperature (293K) experimental lattice constants [4] of a = 4.4284 Å and c = 6.36805 Å. We first performed *ab-initio* calculations for the ionic species, Cr^{2+} and Si^{-}, to generate input orbitals for the solid calculation. Our program package expanded the radial part of the atomic wave functions in terms of Gaussian functions by utilizing a set of even-tempered Gaussian exponents. For Cr^{2+}, our computations utilized 18, 18 and 16 even-tempered Gaussian exponents for the s, p, and d orbitals, respectively. For Si^{1-}, we utilized 18, 18 and 16 even-tempered Gaussian exponents for the s, p, and d orbitals, respectively. Our maximum exponent for Cr^{2+} is 1.1×10^5, while the minimum exponent is 0.317. Similarly, our maximum exponent for Si^{-} is 9.85×10^5, while the minimum exponent is 0.4045. Our computations utilized a mesh of 24 k-points in the irreducible Brillouin zone. However, in the band structure calculation, we utilized a total of 141 weighted k-points while a total of 144 weighted k-points was used in generating the energy eigenvalues for the electronic density of states. Self-consistency was reached after 60 iterations; then, the difference in potentials from any two consecutive calculation was equal to (or less than) 10^{-5}.

In the next section, we present results from our calculation of the band structure, density of states (DOS) and partial density of states (pDOS), and hole effective masses, respectively, using the LDA BZW-EF method.

3. Results

We list below, in Table 4, the valence orbitals in the successive calculations described above, along with the resulting band gaps. The orbitals in bold are the

Table 4. Successive, self-consistent calculations for $CrSi_2$, along with the valence orbitals and the resulting, indirect band gaps. The utilized room temperature lattice constants are a = 4.4284 Å and c = 6.36805 Å. Calculation IV, whose number is in bold in the first column, provided the DFT description of the material, with the corresponding, calculated, indirect band gap of 0.313 eV.

No	Chromium ($3Cr^{2+}$) ($1s^2$ $2s^2$ $2p^6$ ~ Core)	Silicon ($6Si^{1-}$) ($1s^2$ ~ Core)	No of Valence Functions	Indirect Energy Gap (eV) [L – M]
I	$3s^2$ $3p^6$ $3d^4$ **4p**	$2s^2$ $2p^6$ $3s^2$ $3p^3$	168	0.121
II	$3s^2$ $3p^6$ $3d^4$ **4p**	$2s^2$ $2p^6$ $3s^2$ $3p^3$ **4p**	204	0.162
III	$3s^2$ $3p^6$ $3d^4$ **4p 4d**	$2s^2$ $2p^6$ $3s^2$ $3p^3$ **4p**	234	0.295
IV	$3s^2$ $3p^6$ $3d^4$ **4p 4d 4s**	$2s^2$ $2p^6$ $3s^2$ $3p^3$ **4p**	**240**	**0.313**
V	$3s^2$ $3p^6$ $3d^4$ **4p 4d 4s**	$2s^2$ $2p^6$ $3s^2$ $3p^3$ **4p 4s**	252	0.314
VI	$3s^2$ $3p^6$ $3d^4$ **4p 4d 4s 5p**	$2s^2$ $2p^6$ $3s^2$ $3p^3$ **4p 4s**	270	0.318
VII	$3s^2$ $3p^6$ $3d^4$ **4p 4d 4s 5p**	$2s^2$ $2p^6$ $3s^2$ $3p^3$ **4p 4s 5p**	306	0.310

ones representing excited states. Calculation IV was the first one to produce the minima of the occupied energies; the same occupied energies were obtained with Calculations V and VI, signifying that these minima are the absolute ones and represent the ground state, as opposed to being local minima. Figure 2 shows the electronic energy bands for chromium disilicide, along with the bands from Calculations IV and V. As explained above, the two calculations result in the same occupied energies.

Figure 2 shows the electronic band structure of $CrSi_2$ as obtained with Calculation IV. The same figure shows the band structure from Calculation V. As stated above, the occupied energies from these calculations are identical. However, for conduction band energies above 4 eV, the two band structures are different, as explained in the Section on our distinctive, computational method.

Given the large number of bands immediately below and above the Fermi level, in Figure 2, a clear appreciation of their features is difficult. The magnified bands between −3 eV and +3 eV are shown in Figure 3 that provides a clearer view of the features of the DFT band structure in the vicinity of the Fermi level. In this figure, the valence band maximum (VB_{max}) is clearly at the L point, while the conduction band minimum (CB_{min}) is at the M point. The LDA BZW-EF calculated indirect band gap, from L to M, is 0.313 eV, while the smallest direct band gap, at L, is 0.517 eV. This value is only slightly smaller than the L to H and L to K indirect band gaps of 0.533 eV and 0.537 eV, respectively. The values of these gaps can be simply read in Table 5.

Table 5 lists the eigenvalues between −2.748 and +6.094 eV. We expect its content to be useful in comparisons of future experimental findings with our results. Such findings could include optical transition energies and band widths, among others.

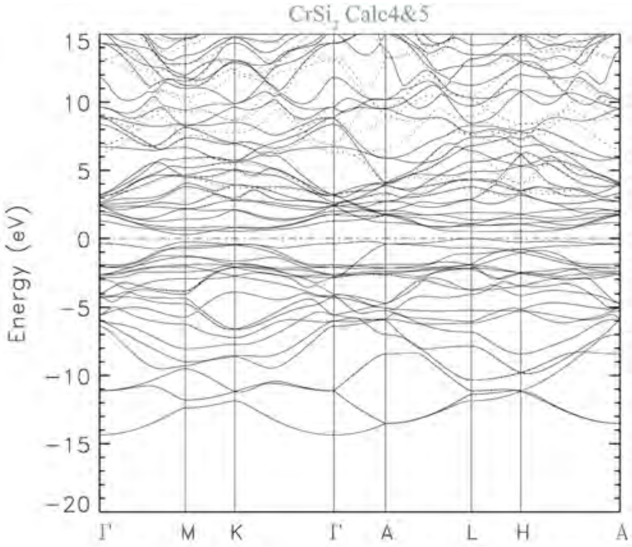


Figure 2. Graphical comparison of Calculations IV and V. Solid line represents Calculation IV while dotted lines represent Calculation V. The Fermi energy level is set at the zero point as denoted by the dashed line at the top of the valence band.

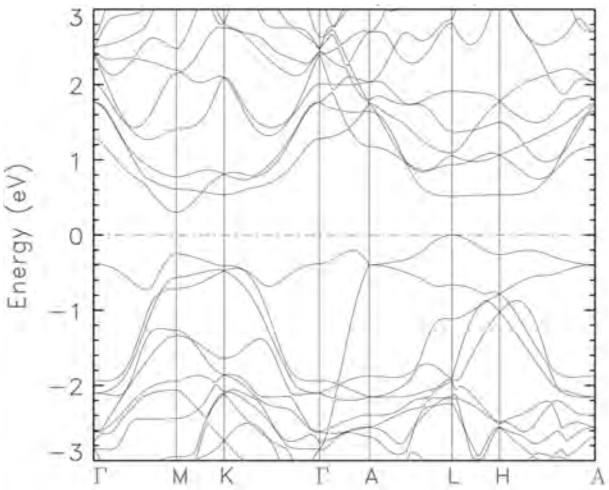


Figure 3. Electronic band structure of $CrSi_2$, as obtained from our ab initio calculations, using ours LDA BZW-EF optimal basis set of Calculation IV.

Table 5. Calculated electronic energies (in eV) of $CrSi_2$, between −2.748 and +6.094 eV, at the high symmetry points in the Brillouin zone, as obtained from Calculation IV. The Fermi energy is set equal to zero. Our calculated indirect band gap is 0.313 eV.

Γ-point	M-point	K-point	A-point	L-point	H-point
3.229	5.197	5.426	4.025	5.625	6.094
3.149	4.666	5.003	3.908	4.833	4.036
3.149	4.568	4.985	3.908	4.364	4.036
2.630	4.470	3.752	3.221	4.270	3.425
2.472	4.066	3.752	2.698	3.712	3.424
2.472	3.793	2.819	2.698	2.866	3.339
2.437	2.477	2.819	2.034	2.821	3.193
2.437	2.194	2.756	2.034	1.914	1.780
2.401	2.150	2.096	1.788	1.365	1.780
2.009	1.411	2.096	1.751	1.090	1.496
1.766	0.774	0.810	1.751	1.056	1.063
1.765	0.609	0.810	1.644	0.934	1.063
1.284	**0.313**	0.537	1.176	0.517	0.533
−0.383	−0.251	−0.411	−0.396	**0.000**	−0.260
−1.934	−0.563	−0.475	−0.397	−0.663	−0.782
−2.103	−0.711	−0.475	−1.874	−1.117	−0.782
−2.103	−1.266	−1.645	−2.155	−1.912	−1.024
−2.615	−1.339	−1.859	−2.155	−1.921	−1.024
−2.616	−1.943	−1.859	−2.394	−1.931	−2.481
−2.748	−2.061	−2.086	−2.557	−2.149	−2.516

The total density (DOS) and partial densities (pDOS) of states, shown in Figure 4 and Figure 5, respectively, provide further insight on the electronic structure. We employed the linear tetrahedron method [66] for the calculations of these densities of states, using the energy bands obtained with the optimal basis set, as shown in Figure 2. The broad peak features of the total density of states reflect the presence of three formula units per primitive cell. While both Cr and Si contribute to this feature between −5 and +5 eV, Si contributions clearly dominate outside this range, as per the partial densities of states. The calculated total width of the valence is 14.38 eV. The inset in Figure 4 shows a detailed view of the boundaries of the band gap.

We have calculated the electron effective masses, in the immediate vicinity of the minimum of the conduction band, at the M point and the hole effective masses, at the maximum of the valence bands, at M. Our calculated electron effective masses along $M\Gamma$, MK, and ML directions are 0.81, 0.77, and 1.38, respectively, in units of free electron mass (m_0). The calculated hole effective masses

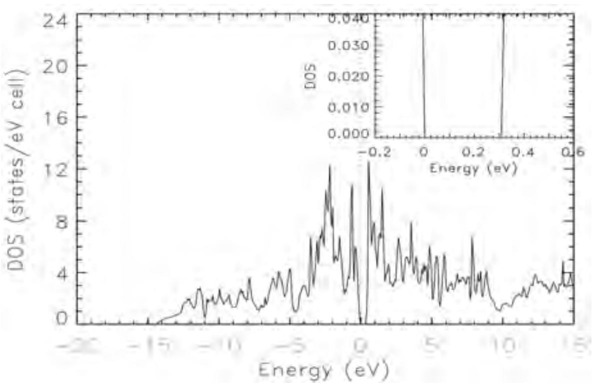


Figure 4. Results from the calculation of the density of states (DOS) for $CrSi_2$, as obtained using the bands from Calculation IV.

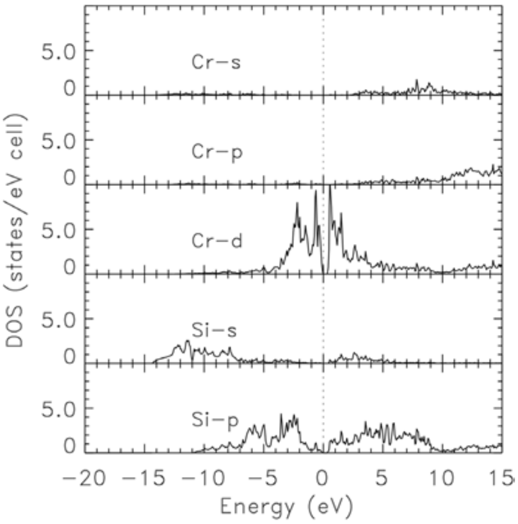


Figure 5. Results from the calculation of partial density of states (pDOS) for $CrSi_2$, as derived from the bands resulting from Calculation IV.

along *LA*, *LH*, *LM*, and *LΓ* axes are 1.3, 1.25, 1.19, and 1.07, respectively, in units of free electron mass. The electron and hole effective masses have been previously calculated by Mattheiss [11] who found that the components of the hole effective mass along *LA*, *LH*, and *LM* axes are 1.2 m_0, 1.3 m_0, 0.9 m_0, respectively. This author also reported electron effective masses of 0.7 m_0, 0.7 m_0, and 1.4 m_0, respectively. While our results for the electron effective masses are only slightly larger than or equal to the corresponding findings of Mattheiss, our hole effective masses, in the LA and LH directions, are much larger than those reported by Mattheiss. Our values somewhat are similar to those found by Mattheiss who used a completely different method (LAPW). Our calculated values for the effective masses are substantially smaller than the corresponding, empirical values of ~3 m_0 and ~20 m_0 for hole and electron effective masses, respectively, as determined from an analysis of transport data [25]. Clearly, more experimental measurements of effective masses in $CrSi_2$ are needed.

4. Discussion

There is a clear need for additional experimental studies of bulk $CrSi_2$. Indeed, as per the content of **Table 1**, only one (1) of the 11 experimental values for the band gap is for bulk $CrSi_2$. The author who reported this value of 3.2 eV did not specify whether the gap was direct or indirect. The other results are indirect band gaps for films of various thicknesses, fabricated by diverse growth techniques. In light of issues of quality of these films and in particular, the well-known quantum confinement effect, which tends to enlarge the gaps of films as compared to bulk materials, there is not much merit in comparing the calculated values for the bulk to these film gaps. The theoretical band gaps in **Table 2** are generally around 3.0 or 3.5 eV, except for the lower value of 0.21 eV and the negative one of −0.35 eV. Even though most of these theoretical results are not too far from the experimental one of 3.2 eV, the fact remains that our finding of 0.313 eV is the closest to this experimental finding. This agreement is partly explained in the Section on our method. Indeed, the BZW-EF method strictly adheres to the conditions of validity of a DFT calculation, *i.e.*, keeping the total number of particles constant and, verifiably, attaining the absolute minima of the occupied energies (the ground state) [53]. The latter condition is imposed by the second DFT theorem. As already noted, this condition is generally far from being met by results from self-consistent iterations with a single basis. A single basis set leads to a stationary solution among an infinite number of them. The relatively better agreement between our calculated band gap and the only experimental one for the bulk stems from the fact that our results possess the full physical content of DFT.

5. Conclusion

We have reported results for the ground state electronic structure and related properties of $CrSi_2$, using the BZW-EF method. Our LDA BZW-EF calculated band gap of 0.313 eV is indirect. Our results for the band gap, total and partial densities of states, and the electrons and hole effective masses are expected to be confirmed by future experimental studies.

Acknowledgements

This research work was funded in part by the US Department of Energy, National and Nuclear Security Administration (NNSA) [Award No. DE-NA0003679], the US National Science Foundation (NSF) [Award No. HRD-1503226], LaSPACE, and LONI-SUBR.

Conflicts of Interest

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

References

[1] Fathauer, R.W., Grunthaner, P.J., Lin, T.L., Chang, K.T., Mazur, J.H. and Jamieson, D.N. (1988) *Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena*, **6**, 708-712. https://doi.org/10.1116/1.584352

[2] Vantomme, A., Nicolet, M.A., Long, R.G., Mahan, J.E. and Pool, F.S. (1993) *Applied Surface Science*, **73**, 146-152. https://doi.org/10.1016/0169-4332(93)90159-9

[3] Heck, C., Kusaka, M., Hirai, M., Iwami, M. and Yokota, Y. (1996) *Thin Solid Films*, **281**, 94-97. https://doi.org/10.1016/0040-6090(96)08583-5

[4] Dasgupta, T., Etourneau, J., Chevalier, B., Matar, S.F. and Umarji, A.M. (2008) *Journal of Applied Physics*, **103**, 113516. https://doi.org/10.1063/1.2917347

[5] White, A.E., Short, K.T. and Eaglesham, D.J. (1990) *Applied Physics Letters*, **56**, 1260-1262. https://doi.org/10.1063/1.103334

[6] Perumal, S., Gorsse, S., Ail, U., Prakasam, M., Chevalier, B. and Umarji, A.M. (2013) *Journal of Materials Science*, **48**, 6018-6024. https://doi.org/10.1007/s10853-013-7398-2

[7] Mason, K. N. (1981) *Progress in Crystal Growth and Characterization*, **2**, 269-307. https://doi.org/10.1016/0146-3535(81)90038-1

[8] Hansen, M. and Anderko, K. (1958) Constitution of Binary Alloys. 2nd Edition, McGraw-Hill, New York, p. 989.

[9] Bost, M.C. and Mahan, J.E. (1988) *Journal of Applied Physics*, **63**, 839-844. https://doi.org/10.1063/1.340078

[10] Krivosheeva, A.V., Shaposhnikov, V.L., Krivosheev, A.E., Filonov, A.B. and Borisenko, V.E. (2003) *Semiconductors*, **37**, 384-389. https://doi.org/10.1134/1.1568455

[11] Mattheiss, L.F. (1991) *Physical Review B*, **43**, 12549. https://doi.org/10.1103/PhysRevB.43.12549

[12] Zhou, S., Xie, Q., Yan, W. and Chen, Q. (2009) *Science in China Series G: Physics, Mechanics and Astronomy*, **52**, 46-51. https://doi.org/10.1007/s11433-009-0003-7

[13] Wetzel, P., Pirri, C., Peruchetti, J.C., Bolmont, D. and Gewinner, G. (1988) *Solid State Communications*, **65**, 1217-1220. https://doi.org/10.1016/0038-1098(88)90926-X

[14] Galkin, N.G. (2007) *Thin Solid Films*, **515**, 8179-8188.

https://doi.org/10.1016/j.tsf.2007.02.043

[15] Luches, A., Mulenko, S.A., Veiko, V.P., Caricato, A.P., Chuiko, V.A., Kudryavtsev, Y.V., Valerini, D., *et al.* (2007) *Applied Surface Science*, **253**, 6512-6516. https://doi.org/10.1016/j.apsusc.2007.01.023

[16] Krivosheeva, A.V., Shaposhnikov, V.L. and Borisenko, V.E. (2003) *Materials Science and Engineering*: *B*, **101**, 309-312. https://doi.org/10.1016/S0921-5107(02)00644-X

[17] Shiau, F.Y., Cheng, H.C. and Chen, L.J. (1986) *Journal of Applied Physics*, **59**, 2784-2787. https://doi.org/10.1063/1.336990

[18] Lange, H. (1997) *Physica Status Solidi* (*B*), **201**, 3-65.

[19] Halilov, S.V. and Kulatov, E.T. (1992) *Semiconductor Science and Technology*, **7**, 368. https://doi.org/10.1088/0268-1242/7/3/016

[20] Krijn, M.P.C.M. and Eppenga, R. (1991) *Physical Review B*, **44**, 9042. https://doi.org/10.1103/PhysRevB.44.9042

[21] Wolf, W., Bihlmayer, G. and Blügel, S. (1997) *Physical Review B*, **55**, 6918. https://doi.org/10.1103/PhysRevB.55.6918

[22] Zhou, F., Szczech, J., Pettes, M.T., Moore, A.L., Jin, S. and Shi, L. (2007) *Nano Letters*, **7**, 1649-1654. https://doi.org/10.1021/nl0706143

[23] Henrion, W., Lange, H., Jahne, E. and Giehler, M. (1993) *Applied Surface Science*, **70-71**, 569-572. https://doi.org/10.1016/0169-4332(93)90581-U

[24] Zhu, H.N. and Liu, B.X. (2000) *Applied Surface Science*, **161**, 240-248. https://doi.org/10.1016/S0169-4332(00)00303-2

[25] Nishida, I. and Sakata, T. (1978) *Journal of Physics and Chemistry of Solids*, **39**, 499-505. https://doi.org/10.1016/0022-3697(78)90026-4

[26] Shinoda, D., Asanabe, S. and Sasaki, Y. (1964) *Journal of the Physical Society of Japan*, **19**, 269-272. https://doi.org/10.1143/JPSJ.19.269

[27] Nishida, I. (1972) *Journal of Materials Science*, **7**, 1119-1124. https://doi.org/10.1007/BF00550193

[28] Bellani, V., Guizzetti, G., Marabelli, F., Piaggi, A., Borghesi, A., Nava, F., *et al.* (1992) *Physical Review B*, **46**, 9380. https://doi.org/10.1103/PhysRevB.46.9380

[29] Galkin, N.G., Maslov, A.M. and Konchenko, A.V. (1997) *Thin Solid Films*, **311**, 230-238. https://doi.org/10.1016/S0040-6090(97)00678-0

[30] Kim, K.H., Kang, J.S., Choi, C.K., Lee, J. and Olson, C.G. (1999) *Applied Surface Science*, **150**, 8-12. https://doi.org/10.1016/S0169-4332(99)00251-2

[31] Shugani, M., Aynyas, M. and Sanyal, S.P. (2015) *Journal of Experimental and Theoretical Physics*, **121**, 104-109. https://doi.org/10.1134/S1063776115060151

[32] Lange, H., Giehler, M., Henrion, W., Fenske, F., Sieber, I. and Oertel, G. (1992) *Physica Status Solidi* (*B*), **171**, 63-76.

[33] Williams, A.R., Kübler, J. and Gelatt Jr., C.D. (1979) *Physical Review B*, **19**, 6094. https://doi.org/10.1103/PhysRevB.19.6094

[34] Eyert, V. (2007) The Augmented Spherical Wave Method. In: *Lecture Notes in Physics*, Springer, Berlin Heidelberg.

[35] Perdew, J.P., Burke, K. and Ernzerhof, M. (1996) *Physical Review Letters*, **77**, 3865. https://doi.org/10.1103/PhysRevLett.77.3865

[36] Filonov, A.B., Tralle, I.E., Dorozhkin, N.N., Migas, D.B., Shaposhnikov, V.L., Petrov, G.V., *et al.* (1994) *Physica Status Solidi* (*B*), **186**, 209-215.

[37] Shaposhnikov, V.L., Krivosheeva, A.V., Krivosheev, A.E., Filonov, A.B. and Borisenko, V.E. (2002) *Microelectronic Engineering*, **64**, 219-223. https://doi.org/10.1016/S0167-9317(02)00788-8

[38] Mattheiss, L.F. (1991) *Physical Review B*, **43**, 1863. https://doi.org/10.1103/PhysRevB.43.1863

[39] Mattheiss, L.F. (1992) *Physical Review B*, **45**, 3252. https://doi.org/10.1103/PhysRevB.45.3252

[40] Bhamu, K.C., Sahariya, J. and Ahuja, B.L. (2013) *Journal of Physics and Chemistry of Solids*, **74**, 765-771. https://doi.org/10.1016/j.jpcs.2013.01.020

[41] Bagayoko, D., Zhao, G.L., Fan, J.D. and Wang, J.T. (1998) *Journal of Physics: Condensed Matter*, **10**, 5645. https://doi.org/10.1088/0953-8984/10/25/014

[42] Bagayoko, D., Franklin, L. and Zhao, G.L. (2004) *Journal of Applied Physics*, **96**, 4297-4301. https://doi.org/10.1063/1.1790064

[43] Bagayoko, D. and Franklin, L. (2005) *Journal of Applied Physics*, **97**, Article ID: 123708. https://doi.org/10.1063/1.1939069

[44] Jin, H., Zhao, G.L. and Bagayoko, D. (2007) *Journal of Applied Physics*, **101**, Article ID: 033123. https://doi.org/10.1063/1.2435802

[45] Stewart, A.D. (2015) *Journal of Advances in Physics*, **9**, 2277-2286. https://doi.org/10.24297/jap.v9i1.1463

[46] Franklin, L., Ekuma, C.E., Zhao, G.L. and Bagayoko, D. (2013) *Journal of Physics and Chemistry of Solids*, **74**, 729-736. https://doi.org/10.1016/j.jpcs.2013.01.013

[47] Ekuma, E.C., Franklin, L., Zhao, G.L., Wang, J.T. and Bagayoko, D. (2011) *Canadian Journal of Physics*, **89**, 319-324. https://doi.org/10.1139/P11-023

[48] Ekuma, C.E. and Bagayoko, D. (2011) *Japanese Journal of Applied Physics*, **50**, Article ID: 101103. https://doi.org/10.7567/JJAP.50.101103

[49] Ekuma, C.E., Jarrell, M., Moreno, J. and Bagayoko, D. (2012) *AIP Advances*, **2**, Article ID: 012189.

[50] Mbolle, A., Banjara, D., Malozovsky, Y., Franklin, L. and Bagayoko, D. (2017) *APS Meeting Abstracts*, **62**, 4.

[51] Banjara, D., Malozovsky, Y., Franklin, L. and Bagayoko, D. (2018) *AIP Advances*, **8**, Article ID: 025212. https://doi.org/10.1063/1.4996551

[52] Bamba, C.O., Malozovsky, Y., Franklin, L. and Bagayoko, D. (2016) *APS Meeting Abstracts*, **61**, 2.

[53] Bagayoko, D. (2014) *AIP Advances*, **4**, Article ID: 127104. https://doi.org/10.1063/1.4903408

[54] Nwigboji, I.H., Malozovsky, Y., Franklin, L. and Bagayoko, D. (2016) *Journal of Applied Physics*, **120**, Article ID: 145701. https://doi.org/10.1063/1.4964421

[55] Jin, H., Zhao, G.L. and Bagayoko, D. (2006) *Physical Review B*, **73**, 245214.

[56] Ceperley, D.M. and Alder, B.J. (1980) *Physical Review Letters*, **45**, 566. https://doi.org/10.1103/PhysRevLett.45.566

[57] Vosko, S.H., Wilk, L. and Nusair, M. (1980) *Canadian Journal of Physics*, **58**, 1200-1211. https://doi.org/10.1139/p80-159

[58] Nwigboji, I.H., Ejembi, J.I., Malozovsky, Y., Khamala, B., Franklin, L., Zhao, G. and Bagayoko, D. (2015) *Materials Chemistry and Physics*, **157**, 80-86. https://doi.org/10.1016/j.matchemphys.2015.03.019

[59] Diakite, Y.I., Traore, S.D., Malozovsky, Y., Khamala, B., Franklin, L. and Bagayoko,

D. (2017) *Journal of Modern Physics*, **8**, 531. https://doi.org/10.4236/jmp.2017.84035

[60] Zhao, G.L. and Bagayoko, D. (2000) *New Journal of Physics*, **2**, 16. https://doi.org/10.1088/1367-2630/2/1/316

[61] Bhandari, U., Bamba, C.O., Malozovsky, Y., Franklin, L.S. and Bagayoko, D. (2018) *Journal of Modern Physics*, **9**, 1773-1784. https://doi.org/10.4236/jmp.2018.99111

[62] Bagayoko, D. (2009) A Solution to the Band Gap Catastrophe: Predictive Calculations of Properties of Semiconductors and of Nuclei. *Proceedings of the First International Seminar on Theoretical Physics & National Development*, **1**, 15-35.

[63] Ekuma, C.E., Jarrell, M., Moreno, J. and Bagayoko, D. (2013) *Physics Letters A*, **377**, 2172-2176. https://doi.org/10.1016/j.physleta.2013.05.043

[64] Harmon, B.N., Weber, W. and Hamann, D.R. (1982) *Physical Review B*, **25**, 1109. https://doi.org/10.1103/PhysRevB.25.1109

[65] Feibelman, P.J., Appelbaum, J.A. and Hamann, D.R. (1979) *Physical Review B*, **20**, 1433. https://doi.org/10.1103/PhysRevB.20.1433

[66] Lehmann, G. and Taut, M. (1972) *Physica Status Solidi* (*B*), **54**, 469-477.

Scientific Research Publishing

Journal of Modern Physics, 2018, 9, 2473-2481
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Lorentz Force from a Current-Carrying Wire on a Charge in Motion under the Assumption of Neutrality in the Symmetrical Frame of Reference

Antonio Šunjerga[1,2], Franjo Sokolić[2], Marcos Rubinstein[3], Farhad Rachidi[1]

[1]Electromagnetic Compatibility Laboratory, EPFL, Switzerland
[2]Department of Physics, Faculty of Science, Split, Croatia
[3]University of Applied Sciences Western Switzerland, Yverdon-les-Bains, Switzerland
Email: antonio.sunjerga@epfl.ch

How to cite this paper: Šunjerga, A., Sokolić, F., Rubinstein, M. and Rachidi, F. (2018) Lorentz Force from a Current-Carrying Wire on a Charge in Motion under the Assumption of Neutrality in the Symmetrical Frame of Reference. *Journal of Modern Physics*, **9**, 2473-2481.
https://doi.org/10.4236/jmp.2018.914159

Received: February 21, 2018
Accepted: December 23, 2018
Published: December 26, 2018

Abstract

It is commonly assumed that a wire conducting an electric current is neutral in the laboratory frame of reference (the rest frame of the lattice of positive ions). Some authors consider that the wire is neutral only in a symmetrical frame of reference, in which the velocities of electrons and protons have equal norm and opposite direction. In this paper, we discuss the Lorentz transformation between different frames of reference in the context of the special theory of relativity for a current-carrying conducting wire and a probe charge in motion with respect to the wire. A simple derivation of the Lorentz force in the laboratory frame of reference for the assumed neutrality in a symmetrical frame of reference is presented. We show that the Lorentz force calculated assuming neutrality in the symmetrical frame of reference and the one assuming neutrality in the laboratory frame of reference differ by a term corresponding to a change in the test charge speed of one half the drift velocity of the electrons.

Keywords

Special Theory of Relativity, Current-Carrying Wire, Neutral Frame, Symmetrical Frame, Lorentz Force, Drift Velocity

1. Introduction

The Lorentz force between a current-carrying wire and a charge in motion in the laboratory frame of reference, where the conductor is at rest, is often expressed as:

$$\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B} \tag{1}$$

where $\boldsymbol{B}$ is the magnetic flux density generated by the wire current, which can be calculated by using Biot-Savart's Law [1]. $\boldsymbol{B}$ depends only on the current magnitude regardless of the physics of the motion of the charge carriers in the wire. In (1), q is the charge and $\boldsymbol{v}$ is its velocity. Equation (1) is correct only under the assumption that the wire is neutral in the laboratory reference frame. Otherwise, a second component of the force due to the electric field should also be considered.

The question of the frame of reference, in which the current-carrying wire is neutral, has been the subject of debate in the past years [2]-[10]. Some authors suggest that a neutral wire corresponds to the rest reference frame of the lattice of positive ions (e.g., [2] [6] [7] [8]), considering that electrons are a free ensemble and, therefore, their distances do not change upon acceleration [3]. Others (e.g., [4] [9]) assume that the distances between electrons are also subject to the Lorentz contraction and, therefore, the wire is neutral only in a symmetrical frame of reference in which both electrons and protons have the same speed but move in opposite directions. Although they are conceptually important, to the best of our knowledge none of these theories has been experimentally proven since the drift velocity of the electrons is small and, hence, the effects are negligible. It is therefore necessary to investigate the mechanisms involved in the transition process from no-current wire to current-carrying wire to answer the question of the neutral frame. A more detailed discussion about the issue of the determination of the reference frame in which the current-carrying wire is neutral can be found in [2] and [3].

In this paper, we will first present a simple derivation for the Lorentz force by assuming the Lorentz contraction of distances between electrons and assuming that the wire neutrality occurs in a symmetrical frame of reference (as in [4] [9]). This assumption leads to a modification of (1). The Lorentz force will depend on the physics of the motion of the charges. In Section 2, we will derive the Lorentz force for the classical example of a charge moving parallel to a current-carrying wire [10] (see **Figure 1** for the geometry of the problem), for a symmetrical frame of reference. In Section 3, we will derive the same force considering a laboratory frame of reference (at rest with respect to the lattice of positive ions), by transforming the force from the symmetrical frame of reference. Conclusions will be given in Section 4.

2. Symmetrical Frame of Reference

Let us first examine the theoretical model from [10]. **Figure 2** shows the problem considering two different reference frames moving with respect to each other at a speed v. The first one (left panel) will be labeled S, and the second (right panel) S'.

Let us imagine that in the symmetrical frame of reference S, we have a straight, infinitely-long wire containing positive and negative charges, characterized by

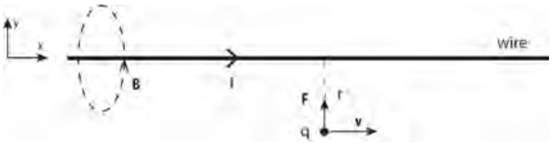


Figure 1. Force between a current-carrying wire and a charge q at a distance r. The charge is moving in the direction parallel to the wire with a speed v.

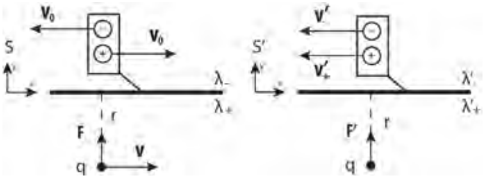


Figure 2. Frame of reference S (left panel) and frame of reference S' (right panel) for $|v| > |v_0|$. Probe charge q. λ_+ and λ_- are the line charge densities in S. λ'_+ and λ'_- are the line charge densities in S'.

linear charge densities λ_+ and λ_-. The positive wire charges move with a constant velocity v_0 in the positive x direction, while the negative wire charges move with a constant velocity-v_0. These two charge densities are measured in the S frame of reference (**Figure 2**, left panel), with respect to which the positive and negative charges are moving. If we assume that both charge densities are equal in absolute value ($|\lambda_+| = |\lambda_-|$), the wire is electrically neutral. Now, let us examine a probe charge q at a distance r moving with a velocity v along a line parallel to the wire in S. In what follows, we will refer to q as the probe charge and to the charges composing the linear charge densities as wire charges.

Let us calculate the force on the probe charge q. Frame S' is moving with a speed v relative to S so that the probe charge q is not in motion in S'. The velocities of the positive and negative line charges in S' can be calculated from the velocities in S by way of the Lorentz transformation:

$$v'_+ = \frac{v_0 - v}{1 - \frac{vv_0}{c^2}}, \quad v'_- = \frac{(-v_0) - v}{1 - \frac{v(-v_0)}{c^2}} \tag{2}$$

Clearly, if v is not equal to zero, v'_+ and v'_- will have different values. Consequently, the distance between the individual positive charges will experience a different contraction compared to the distance between the individual negative charges. We therefore expect different line charge densities λ'_+ and λ'_- in S'. As a result, the wire will not be neutral in S'.

Let us introduce the following parameters [10]:

$$\beta = \frac{v}{c}, \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{3}$$

$$\beta_0 = \frac{v_0}{c}, \quad \gamma_0 = \frac{1}{\sqrt{1-\beta_0^2}} \tag{4}$$

$$\beta'_\pm = \frac{v'_\pm}{c}, \quad \gamma'_\pm = \frac{1}{\sqrt{1-\beta'^2_\pm}} \tag{5}$$

By dividing (2) by c, we can now rewrite β'_+ and β'_- as [10]:

$$\beta'_+ = \frac{\beta_0 - \beta}{1 - \beta_0\beta}, \quad \beta'_- = -\frac{\beta_0 + \beta}{1 + \beta_0\beta} \tag{6}$$

Let us now consider two other frames of reference: 1) the rest frame of reference of the positive wire charges S_+ moving with respect to S with velocity v_0 in the positive x axis direction, and, similarly, 2) the rest frame of reference of the negative wire charges, S_- moving with respect to S with a velocity $-v_0$. Since the linear charge densities are the same in S and since their rest reference frames, S_+ and S_-, are moving with the same speed with respect to S, they experience the same contraction. As a result, the linear charge density of the negative charges in S_- will be the same in magnitude as the linear charge density of the positive charges in S_+. If we label the magnitude of the linear charge density in S as λ ($\lambda = |\lambda_+| = |\lambda_-|$), the linear charge density in the rest frame of reference is simply $\pm\lambda/\gamma_0$, where the sign depends on the wire charge polarity. Now, transforming from S to S', the linear charge density in terms of λ can be expressed as:

$$\lambda'_+ = \gamma'_+ \frac{\lambda}{\gamma_0}, \quad \lambda'_- = -\gamma'_- \frac{\lambda}{\gamma_0} \tag{7}$$

From Equation (7), the overall charge density in S' can be expressed as [10]:

$$\begin{aligned}\lambda'_{overall} &= \lambda'_+ + \lambda'_- = \gamma'_+ \frac{\lambda}{\gamma_0} - \gamma'_- \frac{\lambda}{\gamma_0} = \frac{\lambda}{\gamma_0}\left(\frac{1}{\sqrt{1-\frac{v'^2_+}{c^2}}} - \frac{1}{\sqrt{1-\frac{v'^2_-}{c^2}}}\right) \\ &= \frac{\lambda}{\gamma_0}\left(\frac{1}{\sqrt{1-\beta'^2_+}} - \frac{1}{\sqrt{1-\beta'^2_-}}\right) = -2\lambda\beta\beta_0\gamma = -\frac{2\lambda\gamma v v_0}{c^2}\end{aligned} \tag{8}$$

The electric field from a uniform line charge in S' at the position of the probe charge is given by:

$$E'_y(r) = -\frac{\lambda'_{overall}}{2\pi\epsilon_0 r} \tag{9}$$

As mentioned before, no magnetic force is exerted on the charge since its speed is zero. The total force can therefore be calculated as

$$F'_y(r) = -\frac{q\lambda'_{overall}}{2\pi\epsilon_0 r} \tag{10}$$

Making use of the Lorentz transformation, the force in reference frame S is [10]:

$$F_y(r) = \frac{F'_y}{\gamma} = \frac{q\lambda v v_0}{\pi\epsilon_0 c^2 r} \tag{11}$$

Now, considering that

$$I = \lambda v_0 + (-\lambda)(-v_0) = 2\lambda v_0 \quad \text{and} \quad B(r) = \frac{I}{2\pi\epsilon_0 c^2 r} \tag{12}$$

and plugging these relations into (11), we obtain the more familiar equation:

$$F_y\left(r\right) = qvB \tag{13}$$

This classical example is often used to show the relativistic background of equation (1). In the following section, we will use this derivation to get an expression for the Lorentz Force in the laboratory frame of reference.

3. Laboratory Frame of Reference (Rest Frame of Reference of a Lattice of Positive Charges)

Let us now consider the same case of an infinitely-long current-carrying wire in a laboratory frame of reference. If there is no applied voltage, the random motion of the charges is described with quantum mechanics [11]. If we apply a voltage on the wire, the motion of the electrons will be a superposition of their random motion and that caused by the applied electric field. To a first approximation, the motion can be described as if all electrons were moving with a constant drift velocity [12]. Positive charges are stationary (in lattice) in the considered frame of reference, referred to as the laboratory frame of reference (shown in **Figure 3**). It is assumed that a voltage is applied between the wire ends at infinity, with the left end at a higher potential. The charge q is at a distance r to the wire and it moves with a constant velocity v_{qlab} in the positive x direction, parallel to the wire.

We will now calculate the force applied on the charge in this frame of reference. First, we will transform to the symmetrical frame S (in which the wire is assumed to be neutral) where we will call the speeds of the positive and of the negative charges v_+ and v_- such that:

$$v_+ = -v_- = v_o \tag{14}$$

In this frame of reference, the force is given by Equation (11), in which the value of v, which will be calculated below, is the velocity of the test charge (v_{qlab} in the laboratory frame of reference) with respect to the symmetrical frame of reference. The Lorentz transformation of the positive and negative wire charge velocities to S from $\mathrm{S_{lab}}$ is given by

$$v_+ = -v_S, \quad v_- = \frac{\left(-v_D\right) - v_S}{1 - \frac{v_S\left(-v_D\right)}{c^2}} \tag{15}$$

in which v_s is the velocity of the frame of reference S with respect to $\mathrm{S_{lab}}$.

Substituting Equation (15) in (14), we obtain:

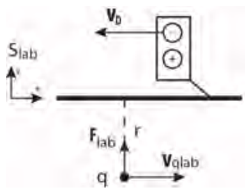


Figure 3. Laboratory frame of reference. Electrons are moving with the drift velocity v_D and positive ions are stationary. The rectangle above the wire represents a differential wire segment illustrating the speeds of the positive and negative charges.

$$\frac{v_D}{c^2}v_S^2 - 2v_S - v_D = 0 \tag{16}$$

Since the drift velocity v_D is in the order of a few mm/s and v_S is even smaller, to a first approximation we can neglect the first term in Equation (16) and obtain the classical Galilean transformation:

$$v_S = -\frac{v_D}{2} \tag{17}$$

In the symmetrical reference frame S, the velocity of the charge is, therefore:

$$v = \frac{v_{qlab} - \left(-\frac{v_D}{2}\right)}{1 - \frac{\left(-\frac{v_D}{2}\right)v_{qlab}}{c^2}} \tag{18}$$

Using Equation (11), the force in the reference frame S becomes:

$$F_y = \frac{q\lambda v\frac{v_D}{2}}{\pi\epsilon_0 c^2 r} \tag{19}$$

In which λ is the magnitude of the positive or the negative charge density in the symmetrical frame of reference in the same way as it was defined in the previous section ($\lambda = |\lambda_+| = |\lambda|$). The Lorentz transformation of this force to the laboratory frame of reference S_{lab} can be expressed as:

$$F_{ylab} = \frac{F_y}{\gamma_S\left(1 - \frac{v\left(-\frac{v_D}{2}\right)}{c^2}\right)} = \frac{q\lambda v\frac{v_D}{2}}{\gamma_S\left(1 + \frac{v\frac{v_D}{2}}{c^2}\right)\pi\epsilon_0 c^2 r} \tag{20}$$

where:

$$\gamma_S = \frac{1}{\sqrt{1 - \frac{v_s^2}{c^2}}} \tag{21}$$

The negative charge density in S_{lab} is:

$$\lambda_{lab-} = -\frac{\lambda}{\gamma_S}\gamma_x \tag{22}$$

in which

$$\gamma_x = \frac{1}{\sqrt{1 - \left(\frac{v_D}{c}\right)^2}} \tag{23}$$

Let us now define:

$$\alpha = \frac{1}{1 + \frac{\frac{v_D}{2}v_{qlab}}{c^2}} \quad \text{and} \quad \delta = \frac{1}{1 + \frac{v\frac{v_D}{2}}{c^2}} \tag{24}$$

Plugging (18) into (20), expressing λ in terms of λ_{lab-} from (20), and using (23) and (24), the Lorentz force in S_{lab} is:

$$F_{ylab}=\frac{q\gamma_S\left(-\lambda_{lab-}\right)\left(\frac{v_{qlab}+\frac{v_D}{2}}{\alpha}\right)\frac{v_D}{2}}{\gamma_S\gamma_x\delta\pi\epsilon_0c^2r}=\frac{q\left(v_{qlab}+\frac{v_D}{2}\right)I}{2\gamma_x\delta\alpha\pi\epsilon_0c^2r}=\frac{q\left(v_{qlab}+\frac{v_D}{2}\right)B}{\gamma_x\delta\alpha} \tag{25}$$

where B and I are defined for the laboratory frame of reference, and where the current is only due to the motion of the negative charges.

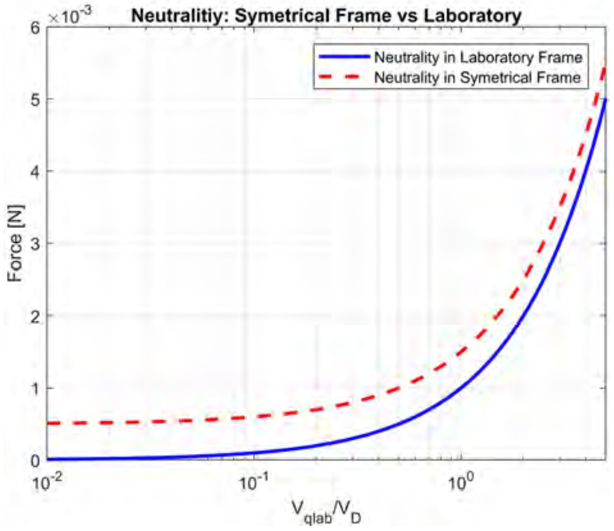
Since $v_D \ll c$, the relativistic coefficient in Equations (23) and (24) is approximately equal to one and (25) can be written as:

$$F_{ylab}\approx q\left(v_{qlab}+\frac{v_D}{2}\right)B \tag{26}$$

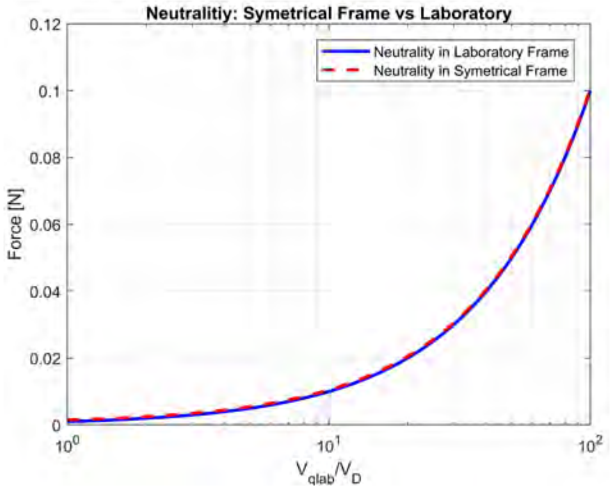
In the laboratory frame of reference (rest frame of the lattice of positive charges), the negative wire charges are moving, as opposed to the positive charges that are stationary. Therefore, they have a higher linear charge density than the stationary positive charges. The overall wire is negatively charged and it produces an electric force in the direction of the magnetic force in the examined case. If the probe charge is moving in the positive direction of the x axis as in **Figure 3** (opposite to the drift of the electrons), force in (1) is increased by $qv_DB/2$ in (26).

It is worth noting that in opposite case when drift velocity of negative wire charges is positive, the sign of force due to probe charge movement will change due to change of sign of current and consequently of magnetic field. However, contribution of this half drift velocity correction term will remain to be directed toward the wire due to change of sign in (17). This is because wire will again be negatively charged in laboratory frame. There will be also change in denominator of two terms in (24), but these terms can be neglected as in (26).

Figure 4 shows comparison of forces calculated in the laboratory frame of reference assuming neutrality in S_{lab} and in the symmetrical frame of reference S for the case of B = 1 T and q = 1 C. As the charge velocity increases, the difference between the forces becomes negligible. The drift velocity [12] is very small, making these effects hard to measure since they are hidden by some other real-scenario forces, such as the zero-order effect of the electrostatic force created by an image charge inside the conducting wire, the first-order force resulting from the resistive nature of the wire, and second-order forces originating for example from the curvature of the wire [4] [13]. In order to mitigate the dominant effect of the zero-order effect force, one may exploit the fact that this force will decay as $1/r^2$ while (26) exhibits a slower $1/r$ decay.



(b)



(b)

Figure 4. Force applied on the charge calculated assuming neutrality in the laboratory frame of reference given with Equation (26) (dashed red curve) and in the symmetrical S frame of reference given by Equation (1) (solid blue curve). For low values of the v_{qlab}/v_D ratio (a) and for high values (b). For the case of $v_D < 0$ and $v_{qlab} > 0$.

4. Conclusion

In this paper, we presented a derivation of the Lorentz force in the laboratory frame of reference for the case of a metallic, current-carrying wire under the assumption of neutrality in the symmetrical frame of reference. The Lorentz force is a combination of the electric and the magnetic forces and, depending on the physics of motion of the charges, the electric field will also be present. We showed that the Lorentz force calculated assuming neutrality in the symmetrical frame of reference and the one assuming neutrality in the laboratory frame of reference differ by a term corresponding to a change in the test charge speed of one half the drift velocity of the electrons. The derived equations make it in principle possible to experimentally test the hypotheses of neutrality. The drift velocity being usually in the order of mm/s, an accurate measurement of these effects might be, however, very challenging, compared to other forces that are in play.

Acknowledgements

Special thanks to Prof. Ante Bilušić for all the great and enthusiastic lectures. Sincere thanks to the two anonymous reviewers whose constructive comments and suggestions improved this paper.

Conflicts of Interest

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

References

[1] Jackson, J. (1999) Classical Electrodynamics. 3rd Edition, John Wiley & Sons, New York.

[2] Redzic, D.V. (2011) *European Journal of Physics*, **33**, 513-523. https://doi.org/10.1088/0143-0807/33/3/513

[3] Folman, R. (2013) *Journal of Physics*: *Conference Series*, **437**, 012013. https://doi.org/10.1088/1742-6596/437/1/012013

[4] Brill, M.H., Dameron, D. and Phipps Jr., T.E. (2011) *Physics Essays*, **24**, 325-326. https://doi.org/10.4006/1.3593402

[5] Redzic, D.V. (2014) *Serbian Astronomical Journal*, No. 188, 55-65. https://doi.org/10.2298/saj1488055r

[6] Feynman, R.P., Leighton, R.B. and Sands, M.L. (2010) The Feynman Lectures on Physics. Volume II New Millenium Edition, New York.

[7] van Kampen, P. (2008) *European Journal of Physics*, **29**, 879-883. https://doi.org/10.1088/0143-0807/29/5/002

[8] Gabuzda, D.C. (1993) *American Journal of Physics*, **61**, 360-362. https://doi.org/10.1119/1.17271

[9] Peters, P.C. (1985) *American Journal of Physics*, **53**, 1165-1169. https://doi.org/10.1119/1.14075

[10] Belsley, M. (2013) *Contemporary Physics*, **54**, 261-262. https://doi.org/10.1080/00107514.2013.836250

[11] Kittell, C. (2005) Introduction to Solid State Physics. 8th Edition, John Wiley & Sons, New York.

[12] Halliday, D., Resnick, R. and Walker, J. (2014) Fundamentals of Physics. 10th Edition, John Wiley & Sons, Hoboken, NJ.

[13] Assis, A.K.T. and Hernandes, J.A. (2007) The Electric Force of a Current. C. Roy Keys, Montreal.

Journal of Modern Physics, 2018, 9, 2482-2494
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

To the Schwarzschild Solution in General Relativity

Valery V. Vasiliev, Leonid V. Fedorov

Institute of Problems in Mechanics, Russian Academy of Sciences, Moscow, Russia
Email: vvvas@dol.ru

How to cite this paper: Vasiliev, V.V. and Fedorov, L.V. (2018) To the Schwarzschild Solution in General Relativity. *Journal of Modern Physics*, **9**, 2482-2494.
https://doi.org/10.4236/jmp.2018.914160

Received: November 28, 2018
Accepted: December 24, 2018
Published: December 27, 2018

Abstract

The paper is concerned with the history of the spherically symmetric static problem solution of General Relativity found in 1916 by K. Schwarzschild [1] [2] which is interpreted in modern physics as the background of the objects referred to as Black Holes. First, the modern interpretation this solution which does not exactly coincide with original solution obtained by K. Schwarzschild is discussed. Second, the basic equations of the original Schwarzschild solution are presented in modern notations allowing us to compare existing and original solutions. Finally, a modification of the Schwarzschild approach is proposed allowing us to arrive at the exact solution of the Schwarzschild problem.

Keywords

General Relativity, Spherically Symmetric Problem, Schwarzschild's Solution, Black Holes

1. Spherically Symmetric Static Problem of General Relativity

Spherically symmetric problem is one of the most discussed problems of General Relativity Theory (GRT) widely described in the literature [3] [4] [5] [6] [7]. This paper is concerned with the analysis of the original Schwarzschild solution of this problem in association with its modern interpretation and possible generalization.

The line element for the spherically symmetric problem is traditionally taken in the following form:

$$\mathrm{d}s^2 = g^2\mathrm{d}r^2 + \rho^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right) - h^2c^2\mathrm{d}t^2 \tag{1}$$

Here r,θ,φ and t are space spherical and time coordinates, g,ρ,h are the metric coefficients that depend on the radial coordinate r only. The basic equations of the General Relativity Theory (GRT) link the Einstein tensor E_i^i with the metric coefficients as [3]

$$E_1^1 = \frac{1}{\rho^2} - \frac{1}{g^2}\frac{\rho'}{\rho}\left(\frac{\rho'}{\rho} + \frac{2h'}{h}\right) \tag{2}$$

$$E_2^2 = E_3^3 = -\frac{1}{g^2}\left[\frac{h''}{h} + \frac{\rho''}{\rho} + \frac{\rho'}{\rho}\left(\frac{h'}{h} - \frac{g'}{g}\right) - \frac{g'h'}{gh}\right] \tag{3}$$

$$E_4^4 = \frac{1}{\rho^2} - \frac{1}{g^2}\left[\left(\frac{\rho'}{\rho}\right)^2 + \frac{2\rho''}{\rho} - \frac{2\rho'g'}{\rho g}\right] \tag{4}$$

where $(\ldots)' = \mathrm{d}(\ldots)/\mathrm{d}r$. The components of the Einstein tensor are proportional to the energy tensor T_i^i, *i.e.*

$$E_i^i = \kappa T_i^i \tag{5}$$

where

$$\kappa = 8\pi\gamma/c^4 \tag{6}$$

is the GRT gravitational constant depending on the Newton constant γ and the velocity of light c. Finally, the energy tensor (and the Einstein tensor which is proportional to it) must satisfy the following conservation equation:

$$\left(T_1^1\right)' - \frac{2\rho'}{\rho}\left(T_2^2 - T_1^1\right) + \frac{h'}{h}\left(T_1^1 - T_4^4\right) = 0 \tag{7}$$

We use mixed components of the tensors E and T because for the problem under study they coincide with the corresponding physical components. The energy tensor depends on the space structure. Particularly, for the empty space

$$T_i^i = 0, (i = 1,2,3,4) \tag{8}$$

and Equations (2)-(4) are homogeneous. Inside a solid sphere with radius a,

$$T_1^1 = \sigma_r, T_2^2 = T_3^3 = \sigma_\theta, T_4^4 = \mu c^2 \tag{9}$$

where σ_r and σ_θ are the radial and the circumferential stresses and μ is the material density.

Consider the external space $(r \ge a)$. Taking $E_4^4 = 0$, we can reduce Equation (4) to the following form:

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{\rho_e\left(\rho_e'\right)^2}{g_e^2}\right] = \rho_e' \tag{10}$$

Equation (10) can be readily integrated to give

$$g_e^2 = \frac{\rho_e\left(\rho_e'\right)^2}{\rho_e + C_1} \tag{11}$$

where C_1 is the integration constant and the functions with subscript "*e*" correspond to the external space. Substituting this result in Equation (2), taking $E_1^1 = 0$ and integrating, we get

$$h_e^2 = C_2\left(1+\frac{C_1}{\rho_e}\right) \tag{12}$$

It looks like substituting Equations (11) and (12) into Equation (3) in which $E_2^2 = 0$, we can determine the function $\rho_e(r)$. But this is not the case—under such substitution, Equation (3) is satisfied identically for any function $\rho_e(r)$. This result can be predicted—since the components of the Einstein tensor E_i^i satisfy Equation (7), Equations (2)-(4) are not mutually independent and any solution of two of them identically satisfy the third equation.

Thus, we have only two Einstein equations for three unknown functions $g(r)$, $\rho(r)$ and $h(r)$. The fact that the set of GTR equations is not complete and must be supplemented with some coordinate conditions was first mentioned by D. Hilbert [8]. By now, the general form of these conditions has not been developed, though some particular forms (e.g., the so-called De-Donder-Fock harmonic coordinate conditions) have been used in spherically symmetric problem [9].

Not knowing the function $\rho_e(r)$, we can make some qualitative conclusions about its behavior. It is natural to suppose the for $r \to \infty$ we should have $\rho_e \to r$ and Equations (11) and (12) should reduce to the corresponding results of the Newton gravitation theory, according to which [10]

$$g_\infty^2 = 1-\frac{2\varphi}{c^2}, \quad h_\infty^2 = 1+\frac{2\varphi}{c^2} \tag{13}$$

where $\varphi = -\gamma m/r$ is the Newton gravitational potential and m is the mass inducing the gravitation field. Equations (13) allows us to determine the constants in Equations (11) and (12) and to present the result as

$$g_e^2 = \frac{(\rho_e')^2}{1-r_g/\rho_e}, \quad h_e^2 = 1-\frac{r_g}{\rho_e} \tag{14}$$

Here,

$$r_g = 2m\gamma/c^2 \tag{15}$$

is the so-called gravitational radius sometimes referred to as the Schwarzschild radius (though K. Schwarzschild did not use this term).

Consider the internal space $(0 \le r \le a)$. For an elastic sphere with known density, we have totally four equations, *i.e.*, Equations (2)-(4) in which the left-hand parts are specified by Equations (5) and (9) and Equation (7) in which the energy tensor should be expressed with the aid of Equation (9). These equations include five unknown functions—the metric coefficients g, ρ, h and the stresses σ_r and σ_θ. To solve the problem, we must supplement the GRT equations with the equations for stresses similar to the compatibility equations of the Theory of Elasticity. Such equations can be derived [11], but we restrict ourselves to a particular solution obtained by K. Schwarzschild for a sphere consisting of a perfect incompressible fluid. In this case, $\sigma_r = \sigma_\theta = -p(r)$ in which p is the pressure in the fluid and the fluid density μ_0 does not depend on p and r.

Taking $E_4^4 = \kappa\mu_0 c^2$ in Equation (4), we can reduce it to the form [12]

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{\rho_i}{g_i^2}\left(\rho_i'\right)\right] = \left(1 - \kappa\mu_0 c^2 \rho_i^2\right)\rho_i' \tag{16}$$

Equation (16) can be readily integrated to give

$$g_i^2 = \frac{\left(\rho'\right)^2}{1 - \lambda^2\rho_i^2 - C_3/\rho_i} \tag{17}$$

where

$$\lambda^2 = \frac{1}{3}\kappa\mu_0 c^2 \tag{18}$$

and the functions with subscript "*i*" correspond to the internal space. We do not know the function $\rho_i(r)$, but can make some reasonable prediction concerning its behavior. Particularly, it is natural to suppose that at the sphere center $\rho_i = r = 0$. Then, in accordance with the symmetry condition at the sphere center, we should take $C_3 = 0$ and Equation (17) becomes

$$g_i^2 = \frac{\left(\rho_i'\right)^2}{1 - \lambda^2\rho_i^2} \tag{19}$$

Continue the derivation and consider Equations (2) and (7). Taking $E_1^1 = -\kappa p$ and $T_1^1 = T_2^2 = -p$, $T_4^4 = \mu_0 c^2$, we arrive at

$$\frac{1}{\rho_i^2} - \frac{\rho_i'}{g_i^2\rho_i}\left(\frac{\rho_i'}{\rho_i} - \frac{2h_i'}{h_i}\right) = -\kappa p,\quad p' + \frac{h_i'}{h}\left(p + \mu_0 c^2\right) = 0 \tag{20}$$

Consider the first Equation. Substituting Equation (19), express the time metric coefficient h_i, *i.e.*,

$$\frac{\mathrm{d}h_i}{h_i} = \frac{\mathrm{d}\rho_i}{2}\left[\frac{\rho_i}{1 - \lambda^2\rho_i^2}\left(\kappa p + \frac{1}{\rho_i^2}\right) - \frac{1}{\rho_i}\right] \tag{21}$$

and rewrite the second equation of Equations (20) as

$$\mathrm{d}p + \frac{\mathrm{d}h_i}{h_i}\left(p + \mu_0 c^2\right) = 0 \tag{22}$$

Substituting Equation (21) and using Equation (18), we arrive at the following equation for the pressure:

$$\mathrm{d}p + \frac{\lambda^2\rho_i\mathrm{d}\rho_i}{2\left(1 - \lambda^2\rho_i^2\right)}\left(p + \mu_0 c^2\right)\left(1 + \frac{3p}{\mu_0 c^2}\right) = 0 \tag{23}$$

The solution of Equation (23) that satisfies the boundary condition on the free sphere surface $r = a$ according to which $p\left(\rho_i = \rho_a\right) = 0$ is

$$p = -\mu_0 c^2 \frac{\sqrt{1 - \lambda^2\rho_i^2} - \sqrt{1 - \lambda^2\rho_a^2}}{\sqrt{1 - \lambda^2\rho_i^2} - 3\sqrt{1 - \lambda^2\rho_a^2}} \tag{24}$$

where $\rho_a = \rho_i(r = a) = \rho_e(r = a)$. To determine h_i, integrate Equation (22) to get $h_i\left(p + \mu_0 c^2\right) = C_4$ in which C_4 is the integration constant that can be found from the boundary condition on the sphere surface $r = a$ according to which $h_e(\rho_e = \rho_a) = h_i(\rho_i = \rho_a)$. Using the second equation in Equations (14), we finally have

$$h_i = \frac{\sqrt{1 - r_g/\rho_a}}{2\sqrt{1 - \lambda^2 \rho_a^2}}\left(3\sqrt{1 - \lambda^2 \rho_0^2} - \sqrt{1 - \lambda^2 \rho_i^2}\right) \tag{25}$$

It should be noted that substitution of the obtained solutions $g_i(\rho_i), h_i(\rho_i)$ and $p(\rho_i)$ in the Einstein Equation (3) where $E_2^2 = -\kappa p$, does not allow us to find the function $\rho_i(r)$ because Equation (3) is identically satisfied for any function $\rho_i(r)$. The situation is similar to the external space and for the same reason—since the right-hand parts of Equations (2) and (3) satisfy Equation (7), only three of Equations (2)-(4) and (7) are mutually independent.

Thus, Equation (1) which specifies the metric forms of the external and internal spaces of the fluid sphere can be presented as

$$\mathrm{d}s_e^2 = \frac{(\rho_e')^2 \mathrm{d}r^2}{1 - r_g/\rho_e} + \rho_e^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2\right) - \left(1 - \frac{r_g}{\rho_e}\right)c^2\mathrm{d}t^2 \tag{26}$$

$$\begin{aligned}\mathrm{d}s_i^2 = {} & \frac{(\rho_i^2)\mathrm{d}r^2}{1 - \lambda^2\rho_i^2} + \rho_i^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2\right) \\ & - \frac{1 - r_g/\rho_a}{4\left(1 - \lambda^2\rho_a^2\right)}\left(3\sqrt{1 - \lambda^2\rho_a^2} - \sqrt{1 - \lambda^2\rho_i^2}\right)^2 c^2\mathrm{d}t^2\end{aligned} \tag{27}$$

To fulfill the solution, we need to find two functions $\rho_e(r)$ and $\rho_i(r)$ such that allow us to satisfy the boundary conditions at the sphere surface $r = a$, *i.e.*,

$$\rho_e(a) = \rho_i(a), \quad g_e(a) = g_i(a) \tag{28}$$

However, the equations allowing us to determine these functions are missing in GRT. The same problem exists in the general case—as known, the set of Einstein equations is not complete. In the four-dimensional Riemannian space, this set consists of 10 equations

$$E^{ij} = -\kappa T^{ij} \tag{29}$$

for ten components of the metric tensor g^{ij}. However the Einstein tensor satisfies equations which are analogous to Equations (7). As a result, only six of Equations (29) are mutually independent and to determine the metric tensor we should supplement Equations (29) with four coordinate conditions for g^{ij}. Some authors declare that these conditions cannot be covariant because there forms depend on the particular coordinate frame [3] [8]. Consider some particular cases.

2. Modern Interpretation of the Schwarzschild Solution

Traditional description of the Schwarzschild solution can be found elsewhere [3]. The coordinate condition mentioned in the closure of the previous section is taken in the form

$$\rho_e(r) = \rho_i(r) = r \tag{30}$$

though K. Schwarzschild did not use directly this relationship. Applying Equations (30) to Equations (14) and (19), we can specify the metric coefficients for this case and present the metric form in Equation (1) as

$$\mathrm{d}s_e^2=\frac{\mathrm{d}r^2}{1-r_g/r}+r^2\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right)-\left(1-\frac{r_g}{r}\right)c^2\mathrm{d}t^2 \tag{31}$$

$$\begin{aligned}\mathrm{d}s_i^2=&\frac{\mathrm{d}r^2}{1-\lambda^2r^2}+r^2\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right)\\&-\frac{1-r_g/a}{4\left(1-\lambda^2a^2\right)}\left(3\sqrt{1-\lambda^2a^2}-\sqrt{1-\lambda^2\rho_i^2}\right)^2c^2\mathrm{d}t^2\end{aligned} \tag{32}$$

To fulfill the solution, we need to satisfy the boundary conditions (28) on the sphere surface $r=a$. The first condition is satisfied automatically, whereas the second one yields

$$r_g=\lambda^2a^3 \tag{33}$$

However, the parameters r_g and λ^2 are specified by Equations (15) and (18) and are known. So, Equation (33) cannot be satisfied in the general case and the second boundary condition in Equations (28) is violated. Substituting formally Equations (15) and (18) in Equation (33), we arrive at the following expression:

$$m=m_e=\frac{4}{3}\pi\mu_0a^3 \tag{34}$$

which specifies the mass of a homogeneous solid sphere in the Euclidean space. However, the space in GRT is not Euclidean and the mass of the sphere with the metric coefficients corresponding to Equations (19) is

$$\begin{aligned}m&=4\pi\mu_0\int_0^a g_ir^2\mathrm{d}r=\frac{2}{\lambda^2}\pi a\mu_0\left(\frac{1}{\lambda a}\sin^{-1}\lambda a-\sqrt{1-\lambda^2a^2}\right)\\&\approx m_e\left(1+\frac{15}{128}\lambda^2a^2+\cdots\right)\end{aligned} \tag{35}$$

As can be seen, the obtained result does not coincide with Equation (34) and the second boundary condition in Equations (28) is not satisfied. The reason for this discrepancy is associated with Equations (30). Equation (4), being originally of the second order, under transformation in accordance with Equation (30) reduces to the equation of the first order. As a result, the solution does not contain the proper number of integration constants allowing us to satisfy the complete set of the boundary conditions.

Thus, the coordinate conditions in Equations (30) do not look suitable for the problem under study.

3. Original Schwarzschild's Solution

In 1916 K. Schwarzschild presented the solution of the external spherically symmetric problem [1]. He did not use the final version of the Einstein equations, however the field equations that he applied can be now associated with

Equations (2)-(4). We can suppose that he understood that only two of these equations were mutually independent because he attracted for the analysis only two equations, *i.e.* Equations (2) and (4), and ignored Equation (3). The third equation which is necessary to solve the problem, was obtained under the following condition imposed on the determinant of the metric tensor:

$$\left|g_{mn}\right| = 1 \tag{36}$$

Introducing this equation, K. Schwarzschild followed A. Einstein who used it in general theory to specify the coordinate frame [13]. Governing equations of GRT contain symbols Γ_{ij}^{k} which include derivatives of the determinant and become zero under condition (33) thus simplifying the equations. However, Equation (33) cannot be directly applied in spherical coordinates in which the volume element in the three-dimensional Euclidean space is $\mathrm{d}v = r^2 \sin\theta \mathrm{d}r \mathrm{d}\theta \mathrm{d}\varphi$. To overcome this problem, K. Schwarzschild introduced new variables x_i such that

$$x_1 = r^3/3,\ x_2 = -\cos\theta,\ x_3 = \varphi,\ x_4 = t \tag{37}$$

In new coordinates, the volume element becomes $\mathrm{d}v = \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3$ and the line element takes the form

$$\mathrm{d}s^2 = f_1 \mathrm{d}x_1^2 + f_2\left[\frac{\mathrm{d}x_2^2}{1-x_2^2} + \left(1-x_2^2\right)\mathrm{d}x_3^2\right] - f_4 \mathrm{d}x_4^2 \tag{38}$$

Three functions f_1, f_2, f_4 can be found from Equations (2) and (4) supplemented with equation $f_1 f_2^2 f_4 = 1$ which follows from Equation (36). The final solution is

$$f_1 = \frac{\left(3x_1 + \beta^3\right)^{-4/3}}{1-\alpha\left(3x_1+\beta^3\right)^{-1/3}},\quad f_2 = \left(3x_1+\beta^3\right)^{2/3},\quad f_4 = 1-\alpha\left(3x_1+\beta^3\right)^{-1/3} \tag{39}$$

in which α and β^3 are the integration constants. As stated in the Schwarzschild paper, this solution identically satisfies Equation (3) (which should be the case).

The final part of the paper can hardly be understood. Directly following K. Schwarzschild, consider the function f_1 which can be a source of singularity. Equating the denominator to zero and using Equation (37) for x_1 we get

$$\alpha\left(r^3+\beta^3\right)^{-1/3} = 1 \tag{40}$$

Let singularity take place at the origin $r = 0$. Then, Equation (40) yields $\beta = \alpha$. Introducing the new variable

$$R = \left(r^3+\alpha^3\right)^{1/3} \tag{41}$$

and using Equations (38) and (39), K. Schwarzschild arrived at the following final result of his paper:

$$\mathrm{d}s^2 = \frac{\mathrm{d}R^2}{1-\alpha/R} + R^2\left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2\right) - \left(1-\alpha/R\right)c^2 \mathrm{d}t^2 \tag{42}$$

This form formally coincides with Equation (31), but it should be taken into account that R is not the radial coordinate r. The constant α is declared to depend on the mass located at the origin, but is not found.

As can be seen, the first term in Equation (42) becomes singular if $R=\alpha$ or $r=0$. Thus, the original Schwarzschild solution has only one singular point—$r=0$.

However, it looks like Equation (42) is not correct. To show this, change R to r in Equation (42) with the aid of Equation (41) to get

$$\begin{aligned}\mathrm{d}s^2 = {} & \frac{\left(r^3+\alpha^3\right)^{-4/3} r^4 \mathrm{d}r^2}{1-\alpha\left(r^3+\alpha^3\right)^{-1/3}} + \left(r^3+\alpha^3\right)^{2/3}\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right) \\ & -\left[1-\alpha\left(r^3+\alpha^3\right)^{-1/3}\right]c^2\mathrm{d}t^2\end{aligned} \tag{43}$$

As can be proved, the first term of this equation becomes zero at $r=0$ which cannot be true. The origin of the mistake is in Equation (40) from which it follows that $\beta=\alpha$.

To demonstrate the alternative approach, substitute Equations (39) in Equation (38) and return to spherical coordinates with the aid of Equations (37). The resulting equation is

$$\begin{aligned}\mathrm{d}s^2 = {} & \frac{\left(r^3+\beta^3\right)^{-4/3} r^4 dr^2}{1-\alpha\left(r^3+\beta^3\right)^{-1/3}} + \left(r^3+\beta^3\right)^{2/3}\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right) \\ & -\left[1-\alpha\left(r^3+\beta^3\right)^{-1/3}\right]c^2\mathrm{d}t^2\end{aligned} \tag{44}$$

As can be seen, the first term has the proper behavior at $r=0$ if we take $\beta=0$ (not $\beta=\alpha$ as in the Schwarzschild solution). The resulting expression

$$\mathrm{d}s^2 = \frac{\mathrm{d}r^2}{1-\alpha/r} + r^2\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right) - \left(1-\frac{\alpha}{r}\right)c^2\mathrm{d}t^2 \tag{45}$$

completely corresponds to the modern interpretation of the Schwarzschild solution in Equation (31), if we apply the asymptotic Equations (13) which give $\alpha=r_g$.

Consider the solution of the internal problem that was published by K. Schwarzschild in 1916 [2]. This solution was not supported by A. Einstein [14] because the concept of an incompressible fluid involves infinitely high velocity of the wave in the fluid which does not correspond to the basic GRT concept. However, the solution for compressible fluids does not demonstrate qualitative deviation from the Schwarzschild solution [15] which is discussed below

The method of the solution is the same that for the external problem, *i.e.*, the new variables in Equations (37) are introduced and the field equations are supplemented with Equation (36). Further, another variable χ is introduced in accordance with the following equation:

$$\sin\chi = \lambda\left(r^3+\eta^3\right)^{1/3} \tag{46}$$

where λ is specified by Equation (18) (in the original equation $c=1$) and η^3 is the integration constant. The final original result for the space part of the metric form is

$$\mathrm{d}s_i^2=\frac{1}{\lambda}\left[\mathrm{d}\chi^2+\sin^2\chi\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right)\right]-\left(\frac{3}{2}\cos\chi_a-\frac{1}{2}\cos\chi\right)^2c^2\mathrm{d}t^2 \quad (47)$$

where $\chi_a=\chi(r=a)$.

To discuss the result obtained by K. Schwarzschild, change χ to r using Equation (46). Then, Equation (47) becomes

$$\begin{aligned}\mathrm{d}s_i^2={}&\frac{\left(r^3+\eta^3\right)^{-4/3}r^4\mathrm{d}r^2}{1-\lambda^2\left(r^3+\eta^3\right)^{2/3}}+\left(r^3+\eta^3\right)^{2/3}\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right)\\&-\frac{1}{4}\left[3\sqrt{1-\lambda^2\left(a^3+\eta^3\right)^{2/3}}-\sqrt{1-\lambda^2\left(r^3+\eta^3\right)^{2/3}}\right]^2c^2\mathrm{d}t^2\end{aligned}$$

The first coefficient becomes zero at $r=0$, which cannot be true. To obtain the realistic metric, we must take $\eta=0$ and arrive at the expression

$$\mathrm{d}s_i^2=\frac{\mathrm{d}r^2}{1-\lambda^2r^2}+r^2\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right)-\frac{1}{4}\left(3\sqrt{1-\lambda^2a^2}-\sqrt{1-\lambda^2r^2}\right)^2c^2\mathrm{d}t^2 \quad (48)$$

In the closure of his paper, K. Schwarzschild analyzed the obtained solution. Particularly, the sphere mass was found in the form

$$m=\frac{6\pi}{\kappa\lambda}\left(\chi_a-\frac{1}{2}\sin2\chi_a\right) \quad (49)$$

which coincides with Equation (35) after transformation with the aid of Equation (46). The first term in Equation (48) becomes singular at some critical sphere radius $a_g=1/\lambda$. This radius coincides with the gravitational radius r_g only under the condition imposed by Equation (33). If this equation is valid, the metric form in Equation (48) coincides with Equation (32). But then, the sphere mass is specified by Equation (34) which corresponds to the Euclidean space. However, the mass found by K. Schwarzschild is given by Equation (49) and corresponds to the Riemannian space. Thus, the radius r_g cannot be called the Schwarzschild radius.

Now return to Equation (44) which specifies the Schwarzschild solution for the external space. Taking $r\to\infty$, and performing asymptotic analysis, we can prove that the metric coefficients in Equation (44) reduce to Equations (13) if we take $\alpha=r_g$. Then Equation (44) becomes

$$\begin{aligned}\mathrm{d}s^2={}&\frac{\left(r^3+\beta^3\right)^{-4/3}r^4\mathrm{d}r^2}{1-r_g\left(r^3+\beta^3\right)^{-1/3}}+\left(r^3+\beta^3\right)^{2/3}\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\varphi^2\right)\\&-\left[1-r_g\left(r^3+\beta^3\right)^{-1/3}\right]c^2\mathrm{d}t^2\end{aligned} \quad (50)$$

In contrast to the traditional Equation (31), this equation contains one more integration constant—β. This result looks natural because K. Schwarzschild did not use Equations (30) and, hence, did not reduce the order of Equation (4). Considering the space with a point mass, we took $\beta = 0$ because of the behavior of the first coefficient in Equation (50) at $r = 0$ and reduced Equation (44) to Equation (45). But now we study the external space of a fluid sphere for which $r \geq a$. So, we can try to use this constant to obtain the continuous solution in Equations (48) and (50) at the sphere surface. However, matching equations (48) and (50), we can conclude that the second terms can be continuous only if $\beta = 0$. Thus, the final form of Equation (50) for the external space of a fluid sphere

$$\mathrm{d}s^2 = \frac{\mathrm{d}r^2}{1 - r_g/r} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right) - \left(1 - \frac{r_g}{r}\right)c^2\mathrm{d}t^2 \tag{51}$$

coincides with the traditional Equation (31). The first and the third coefficients of Equations (48) and (51) are continuous at $r = a$ if $r_g = \lambda^2 a^3$ which coincides with Equation (33). Thus, Equation (36) applied by K. Schwarzschild is not the proper coordinate condition and actually gives the same results that the conditions in Equations (30). The original Schwarzschild solution, as well as it modern interpretation, does not provide the solution which satisfies GRT equations and all asymptotic and boundary conditions for a fluid sphere.

4. New Model of Space and Spherically Symmetric Problem

Traditionally GRT is associated with Riemannian geometry which describes the so-called curved space. A three-dimensional curved space can be hardly imagined. This space can be formally embedded into traditional Euclidean space. However such space has six dimensions which do not provide better understanding of the problem. The proposed interpretation of the Riemannian space is based on the following assumptions. First, we assume that the space is not an object of geometry, but is a material substance (ether, physical vacuum or whatever else). Second, we think that the curved space does not exist in reality and the Riemannian geometry is only a mathematical model of a special Euclidean space. This space is not homogeneous and is characterized with so-called space density that is a function of the coordinates to which the space is referred. The space density $d = \mathrm{d}v_R/\mathrm{d}v_E$ is the ratio of the three-dimensional volume elements corresponding to the Riemannian and to the Euclidean geometries in the same coordinate frame [16] [17]. For spherical coordinates and the metric form in Equation (1), we have $\mathrm{d}v_R = g\rho^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi$ and $\mathrm{d}v_E = r^2\sin\theta\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi$, so that $d = g\left(\rho/r\right)^2$. Using Equations (14) and (19), we get for the external and internal spaces

$$d_e = \frac{\rho_e^2 g_e}{r^2} = \frac{\rho_e'\rho_e^2}{r^2\sqrt{1 - r_g/\rho_e}}, \quad d_i = \frac{\rho_i^2 g_i}{r^2} = \frac{\rho_i'\rho_i^2}{r^2\sqrt{1 - \lambda^2\rho_i^2}} \tag{52}$$

The space densities in Equations (52) are characterized with some specific properties. Consider the total density of the external and internal spaces for the fluid sphere with radius a

$$D_e = 4\pi\int_a^\infty d_e r^2 \mathrm{d}r = 4\pi\int_a^\infty F\left(\rho_e,\rho_e'\right)\mathrm{d}r,\ \ F\left(\rho_e,\rho_e'\right) = \frac{\rho_e'\rho_e^2}{\sqrt{1-r_g/\rho_e}} \tag{53}$$

$$D_i = 4\pi\int_0^a d_i r^2 \mathrm{d}r = 4\pi\int_0^a F\left(\rho_i,\rho_i'\right)\mathrm{d}r,\ \ F\left(\rho_i,\rho_i'\right) = \frac{\rho_i'\rho_i^2}{\sqrt{1-\lambda^2\rho_i^2}} \tag{54}$$

Consider the variational equations providing the minimum values of the functional in Equations (53) and (54), *i.e.*,

$$\frac{\partial F}{\partial \rho} - \frac{\partial}{\partial r}\frac{\partial F}{\partial \rho'} = 0 \tag{55}$$

As can be readily proved, Equation (55) is satisfied identically for both functions F in Equations (53) and (54). Thus, the space densities in Equations (52) provide the minimum total density of the space. However, the space density is caused by gravitation and is minimum in the absence of gravitation, *i.e.*, if the space geometry is Euclidean or if $d=1$ which means that the space tends to become homogeneous with respect to the space density d. The condition $d=1$ looks slightly similar to Equation (36) applied by K. Schwarzschild. It has a simple physical and geometrical meaning—gravitation, changing the space geometry, does not affect the space volume.

Taking $d_e=1$ and $d_i=1$ in Equations (52), we arrive at two differential equations for functions $\rho_e(r)$ and $\rho_i(r)$, *i.e.*,

$$\rho_e'\rho_e^2 = r^2\sqrt{1-r_g/\rho_e}\,,\ \ \rho_i'\rho_i^2 = r^2\sqrt{1-\lambda^2\rho_i^2} \tag{56}$$

Consider the second equation. Since $d_i = g_i\rho_i^2/r^2 = 1$, the sphere mass becomes

$$m = 4\pi\mu_0\int_0^a g_i\rho_i^2 \mathrm{d}r = 4\pi\mu_0\int_0^a r^2\mathrm{d}r = \frac{4}{3}\pi\mu_0 a^3$$

and coincides with the Euclidean mass in Equation (34) which means that the condition in Equation (33) is valid and $\lambda^2 = r_g/a^3$. Then, the second equation in Equations (56) reduces to

$$\rho_i'\rho_i^2 = r^2\sqrt{1-r_g\rho_i^2/a^3} \tag{57}$$

The solution of Equation (57) which satisfies the boundary condition $\rho_i(r=0)=0$ is [12]

$$\frac{1}{\sqrt{\bar{r}_g}}\sin^{-1}\left(\bar{\rho}_i\sqrt{\bar{r}_g}\right) - \bar{\rho}_i\sqrt{1-\bar{r}_g\bar{\rho}_i^2} = \frac{2}{3}\bar{r}_g\bar{r}^3 \tag{58}$$

where

$$\bar{r} = r/a,\ \ \bar{\rho} = \rho/a \tag{59}$$

Recall that at the sphere surface $\rho_i(r=a)=\rho_a$. Taking $\bar{r}=1$ and $\bar{\rho}_i = \bar{\rho}_a$ in Equation (58), we get

$$\frac{1}{\sqrt{\bar{r}_g}}\sin^{-1}\left(\bar{\rho}_a\sqrt{\bar{r}_g}\right) - \bar{\rho}_a\sqrt{1-\bar{r}_g\bar{\rho}_a^2} = \frac{2}{3}\bar{r}_g \tag{60}$$

The general solution of the first equation in Equations (56) is [12]

$$\left(\frac{1}{3}\bar{\rho}_e^2+\frac{5}{12}\bar{r}_g\bar{\rho}_e+\frac{5}{8}\bar{r}_g^2\right)\sqrt{\bar{\rho}_e\left(\bar{\rho}_e-\bar{r}_g\right)}+\frac{5}{8}\bar{r}_g^3\ln\left(\sqrt{\bar{\rho}_e}+\sqrt{\bar{\rho}_e-\bar{r}_g}\right)=\frac{1}{3}\bar{r}^3+C_5 \quad (61)$$

The integration constant can be found from the boundary condition on the so here surface according to which $\bar{\rho}_e\left(\bar{r}=1\right)=\bar{\rho}_a$. Then,

$$C_5=\left(\frac{1}{3}\bar{\rho}_a^2+\frac{5}{12}\bar{r}_g\bar{\rho}_a+\frac{5}{8}\bar{r}_g^2\right)\sqrt{\bar{\rho}_a\left(\bar{\rho}_a-\bar{r}_g\right)}+\frac{5}{8}\bar{r}_g^3\ln\left(\sqrt{\bar{\rho}_a}+\sqrt{\bar{\rho}_a-\bar{r}_g}\right)-\frac{1}{3} \quad (62)$$

Thus, the functions $\rho_i\left(r\right)$ and $\rho_e\left(r\right)$ are specified by Equations (58) and (61). The first boundary condition in Equations (28) according to which $\rho_i\left(a\right)=\rho_e\left(a\right)=\rho_a$ is satisfied which follows from Equations (60) and (62). The second boundary condition in Equations (28) according to which $g_i\left(a\right)=g_e\left(a\right)$ is satisfied because the obtained solution is based on the condition $g_i\rho_i^2=g_e\rho_e^2$ from which it follows that if the function $\rho\left(r\right)$ is continuous at $r=a$, the function $g\left(r\right)$ is also continuous. For low levels of gravitation, *i.e.* for $\bar{r}_g\ll 1$, Equations (58) and (61) yield $\rho_i=\rho_e=r$.

Consider Equations (60) and (62) which allow us to satisfy the boundary conditions, *i.e.*, to solve the problem that turned out to be critical for the solutions discussed above. As follows from Equation (62), the solution exists if $\bar{\rho}_a\geq\bar{r}_g$. Otherwise, the solution becomes imaginary. The minimum possible value of $\bar{\rho}_a$ is $\bar{r}_g$. Assume that this minimum value corresponds to the sphere radius a_g. Then, substituting $\bar{\rho}_a=\rho_a/a_g=r_g/a_g$ in Equation (60), we get

$$\sqrt{\frac{a_g}{r_g}}\sin^{-1}\left(\frac{r_g}{a_g}\sqrt{\frac{r_g}{a_g}}\right)-\frac{r_g}{a_g}\sqrt{1-\left(\frac{r_g}{a_g}\right)^3}=\frac{2r_g}{3a_g} \quad (63)$$

The solution of Equation (63) is $a_g=1.115r_g$. Thus, the critical radius is larger than the gravitational radius. For $a=a_g$, the solution is not singular and gives finite values for the metric coefficients. Particularly, for $\rho=\rho_a$ we get $g_e=g_i=1.243$ and $\rho_e=\rho_i=0.8968a$. For $a<a_g$, the solution becomes imaginary which means that GRT is not valid for such high level of gravitation. For the sphere with radius a_g, the escape velocity is equal to the velocity of light and such sphere is invisible [18]. More results concerning the discussed solution can be found elsewhere [12].

5. Conclusion

As follows from the foregoing analysis, the Schwarzschild solution after some minor correction and reconstruction coincides with the traditional [3] interpretation of this solution. Both solutions do not satisfy the boundary condition on the fluid sphere surface for the radial space metric coefficient. A proposed model of the Riemannian space as the Euclidean space of variable density allows us to obtain the solution which satisfies equations GRT and all boundary conditions for the spherically symmetric problem for a fluid sphere. In future, the authors plan to generalize the approach discussed in Section 4 to the axially symmetric problem of GRT.

Conflicts of Interest

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

References

[1] Schwarzschild, K. (1916) *Sitz. Preuss. Akad. Wiss.*, **1916**, 189-207.

[2] Schwarzschild, K. (1916) *Sitz. Preuss. Akad. Wiss.*, **1916**, 424-432.

[3] Synge, J.L. (1960) Relativity: the General Theory. Amsterdam, North Holland.

[4] Weinberg, S. (1972) Gravitation and Cosmology. Amsterdam, North Holland.

[5] Misner, S.W., Thorne, K.S. and Wheeler, G.A. (1973) Gravitation. W.H. Freeman and Co., San Francisco.

[6] Chandrasecar, S. (1983) The Mathematical Theory of Black Holes. Oxford University Press, Oxford.

[7] Frolov, V.P. and Zelnikov, A. (2011) Introduction to Black Hole Physics. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780199692293.001.0001

[8] Logunov, A.A., Mestvirishvili, M.A. and Petrov, V.A. (2004) *Physics-Uspekhi*, **47**, 607-621. (In Russian) https://doi.org/10.1070/PU2004v047n06ABEH001817

[9] Fock, V. (1959) The Theory of Space, Time and Gravitation. Pergamon Press, London.

[10] Landau, L.D. and Lifshits, E.M. (1988) Field Theory. Nauka, Moscow. (In Russian)

[11] Vasiliev, V.V. and Fedorov, L.V. (2017) *Applied Physics Research*, **6**, 8-13. https://doi.org/10.5539/apr.v9n6p8

[12] Vasiliev, V.V. (2017) *Journal of Modern Physics*, **8**, 1087-1100. https://doi.org/10.4236/jmp.2017.87070

[13] Einstein, A. (1916) *Annalen der Physik*, **49**, 769-819. https://doi.org/10.1002/andp.19163540702

[14] Einstein, A. (1939) *Annals of Mathematics*, **40**, 922-936. https://doi.org/10.2307/1968902

[15] Vasiliev, V.V. and Fedorov, L.V. (2012) *Applied Physics Research*, **2**, 166-174.

[16] Vasiliev, V.V. (1989) *Mechanics of Solids*, **5**, 30-34.

[17] Vasiliev, V.V. and Fedorov, L.V. (2013) *Applied Physics Research*, **1**, 115-125.

[18] Vasiliev, V.V. and Fedorov, L.V. (2015) *Applied Physics Research*, **3**, 84-92.

Scientific Research Publishing

Journal of Modern Physics, 2018, 9, 2495-2593
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Evolution, Quantization, Relativity: An "*Ab Initio*" Model

Sebastiano Tosto[*]

ENEA, via Anguillarese 301, Roma, Italy
Email: stosto44@gmail.com, stosto@inwind.it

How to cite this paper: Tosto, S. (2018) Evolution, Quantization, Relativity: An "*Ab Initio*" Model. *Journal of Modern Physics*, **9**, 2495-2593.
https://doi.org/10.4236/jmp.2018.914161

Received: September 28, 2018
Accepted: December 24, 2018
Published: December 27, 2018

Abstract

"$\pi\alpha'\nu\tau\alpha \quad \chi\omega\rho\varepsilon\tilde{\iota}$ = everything flows", Eraclitus, (Ephesus, 535-475 B.C.). If really in Nature everything changes and progresses, then at least two questions arise: 1) how can be these changes entropic but nonetheless somehow predictable without risk of oxymoronic behavior; 2) how can Science conform itself to follow this requirement of the Nature. To attempt an answer to these questions, the present paper introduces an *ab initio* theoretical model aimed to show that physical information is actually nothing else but straightforward quantum and relativistic implication of the concept of evolution.

Keywords

Evolution, Quantum Theory, Relativistic Theory

1. Introduction

Many physicists have emphasized the unreasonable effectiveness of mathematics in describing the physical world; among them the most authoritative one is Wigner [1]. An anecdote clarifies Wigner's perplexity. Two students were discussing the ability of describing the statistical distribution of hungers in the world through the Gauss function, which involves the number π. Strictly speaking, it is hard to realize what has to do the geometrical ratio between circumference and radius of a circle with the distribution of hungers; even in lack of a rational explanation, though, nobody could doubt about the ability of scientists to contribute to the advancement of science introducing π in the frame of sophisticated mathematical algorithms.

Even Bertrand Russel was concerned about the link between mathematics and physics [2]. In his book "Study of Mathematics" he says: "Mathematics, rightly

---

[*] Retired Physicist.

viewed, possesses not only truth, but supreme beauty, a beauty cold and austere like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show. The true spirit of delight, the exaltation, the sense of being more than Man, which is the touchstone of the highest excellence, is to be found in mathematics as surely as in poetry."

Nevertheless, the outcomes of the natural sciences are subjected to experimental tests: what is false or true is definable regardless of hungers and geometrical distresses. On the one hand, abstract numbers express reliable physical laws describing properties and predicting behavior of Nature. On the other hand, however, this epistemological shortcut in fact leaves unexplained the link between science and reality, calculation and experiment, mental ideas and actual story of the Universe. Quoting Einstein "the most incomprehensible thing of the Universe is that it is comprehensible".

Paradoxically, it is easy to understand the correlation between mathematical algorithm and natural event assuming first deterministic evolution of systems according to the old classical physics: once having selected properly the initial conditions, the successive evolution is in principle uniquely determinable. In practice any deterministic model requires a suitable number of descriptive parameters exactly known of a whole system, whose time evolution is codified and described via appropriate functions of these parameters; the mathematical definitions valid at a given time $t = t_0$, remain also valid, if correctly chosen, at $t_0 + \delta t$. Extrapolating this reasoning, the outcomes of such a model hold at any times $t_0 + n\delta t$ even for $n \to \infty$: everything exactly known at $t = t_0$ remains exactly knowable forever. This should be true in principle also for a classical Universe, regarded as a whole physical system.

Actually however the problem is much more complicated.

The task of guessing the evolution of a physical system from a given initial condition must settle up with the probabilistic frame of the quantum theory: uncertainty relationships imply the impossibility of knowing simultaneously couples of conjugate dynamical variables. This constrain at the time t_0 prevents the possibility of their exact knowledge at any later time as well. Worse still, an initial energy uncertainty $\delta\varepsilon_0$ compels a subsequent range $\delta\varepsilon$ of possible values that depends itself upon the choice of δt. As a matter of fact, however, the fundamental laws of quantum physics are successful in conceiving correctly and designing operatively transistors and lasers.

The predictive ability of science becomes further at stake considering also the relativistic theory, according which δt and $\delta\varepsilon$ have meaning only relatively to the particular reference system where they are initially defined: e.g. the twin paradox exemplifies that the time is not an absolute parameter, as the reciprocal motion of their reference systems R and R' implies anyway admitting different time lapses δt and $\delta t'$ for a given event to occur and even for their aging. The necessity of specifying both reference systems to describe physical events explains why the physical laws must be formulated in a covariant way.

Moreover the link between quantum and relativistic theory is still a hard challenge even today.

To approach gradually the epistemological problem raised by these short considerations, suppose preliminarily that a given event K is allowed to occur in a given R at the arbitrary time t_1, waiving for the moment whether or not actually this time is exactly determinable; in this R are also defined the initial time t_o and the pertinent boundary conditions. Let K be for example the motion of a classical system of N particles, described by a total number J of descriptive parameters f_j: e.g. position coordinates $x_{r,i} = x_{r,i}\left(t_1\right)$, momentum components $p_{r,i} = p_{r,i}\left(t_1\right)$, energies $\varepsilon_i = \varepsilon_i\left(t_1\right)$ and so on of each *i*-th particle with $r = 1,2,3$. Of course f_j can include also mutual interactions, presence of external fields and anything else. Shortly, $f_{t_1,j}$ symbolize in general the *j*-th dynamical variables significant to define the state of each particle at the time t_1, in principle all measurable. Moreover let be known also the experimental value of the observable $V_1 = V\left(t_1\right)$ of a given property V characterizing the event K at the time t_1 and reproduce this value as a linear combination $V_1 = \sum_{j=1}^{J} A_j f_{t_1,j}$ of its descriptive parameters via appropriate coefficients A_j. A simple example clarifies this point. Consider a one dimensional system of two interacting particles having initial coordinates x_{o1} and x_{o2} at the time t_o; concerning first the initial boundary condition, write

$$V_o = A_1 x_{o,1} + A_2 x_{o,2} + A_3 \varepsilon_{o,1} + A_4 \varepsilon_{o,2} + A_5 \varepsilon_o^* + A_6 t_o$$

being all dynamical variables known by definition, regard the coefficients A_j as parameters that fit the initial value V_o of total energy of the system; also, are experimentally measured the space and time coordinates and the energies ε_{o1} and ε_{o2}, upon which depends the interaction energy ε_o^* assumed known as well. Repeat this reasoning, but considering now the total energy V_1 of the system at the arbitrary time t_1; it is in principle possible that the same equation links V_1 to the new space and time coordinates $x_{1,1}$ and $x_{1,2}$, energies $\varepsilon_{1,1}$ and $\varepsilon_{1,2}$, and interaction energy ε_1^*. Assuming experimentally known all these quantities, the new linear combination involving the same dynamical variables reasonably determining V_1 experimentally known as well reads

$$V_1 = A_1 x_{1,1} + A_2 x_{1,2} + A_3 \varepsilon_{1,1} + A_4 \varepsilon_{1,2} + A_5 \varepsilon_1^* + A_6 t_1.$$

In principle it is possible that the coefficients A_j, with $1 \geq j \geq 6$ in this example, fit not only the initial boundary condition but also this further equation. In practice, however, neither the former equation nor the latter are calculable because two equations do not define uniquely the six unknowns A_j; the system of equations is actually undetermined. But it is possible to measure all dynamical variables also at the subsequent arbitrary times t_2 or t_3 and so on, which yields

$$V_k = A_1 x_{1,k} + A_2 x_{2,k} + A_3 \varepsilon_{1,k} + A_4 \varepsilon_{1,k} + A_5 \varepsilon_k^* + A_6 t_k.$$

It is clear that further sets of six experimental data obtained measuring the same quantities at five additional times with respect to the initial condition, yield a system of six equations with six unknowns. Now the system admits a unique solution for all A_j fulfilling by definition also the boundary condition. In principle this empirical procedure is possible no matter how complex is the system and how many its freedom degrees might be; a suitable number J of experimental measurements allow to obtain coefficients A_j that fit by definition all values of the observable V_k of interest at any time between t_o and t_J. The various $f_{t_k,j}$ are therefore not only the respective x_k but also any other dynamical variable that concur with all ε_k and ε_k^* to the resulting value of the observable V_k; it is clear why one of the dynamical variables must be just the pertinent time t_k. Note that, owing to the empirical character of the linear combination, even the higher powers of some descriptive parameters, e.g. $p_{r,i}^2$ are in principle admissible with their own A_j among the terms contributing to V_k.

Anyway take for granted that, by definition, all coefficients A_j fit correctly the known values of the experimental parameters $f_{t_k,j}$ of all particles concurring to the required value V_k.

On the one hand, is comprehensible the interest to describe the system at subsequent times after that of the initial condition for completeness of information. On the other hand, however, since in general the descriptive parameters are functions of time, e.g. the dynamical variables of the various particles, the evolution of the system during a given time range becomes in fact essential requirement for the mathematical approach: repeating the same numerical procedure at $J-1$ subsequent times t_k after the initial t_o, one can define a set of J equations and thus a square matrix of coefficients A_j whose lines fit exactly by definition the experimental values V_k of the observable V in the given time range. Write therefore

$$V_k = \sum_{j=1}^{J} A_j f_{t_k,j},\ 1 \le k \le J,\ V_k = V_k\left(x_j, t_k\right) \tag{1.1}$$

the system of equations removes the indeterminacy inherent a unique observation time and contextually describes how a given observable of the system changes at various times $t_1 \le t_k \le t_J$, although without rational or heuristic valence. Nonetheless the following evolution matrix represents the minimal condition able to characterize mathematically one property V of one event K of the system, although waiving any chance of physical explanation:

$$\boldsymbol{KA} = |\boldsymbol{V}|,\quad \boldsymbol{K} = \begin{pmatrix} f_{t_o,1} & f_{t_o,2} & \cdots & f_{t_o,J} \\ \vdots & \vdots & \ddots & \vdots \\ f_{t_k,1} & f_{t_k,2} & \cdots & f_{t_k,J} \\ \vdots & \vdots & \ddots & \vdots \\ f_{t_J,1} & f_{t_J,2} & \cdots & f_{t_J,J} \end{pmatrix},\quad \boldsymbol{A} = \begin{pmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_J \end{pmatrix},\quad \boldsymbol{V} = \begin{pmatrix} V_{t_o} \\ \vdots \\ V_{t_k} \\ \vdots \\ V_{t_J} \end{pmatrix}, \tag{1.2}$$

Every column of the matrix $\boldsymbol{K}$ represents the values of each descriptive parameter governing V at various times, every line concerns the values of all possible descriptive parameters contributing to the value V_k at the particular time

t_k regardless of explaining how it was at the past t_{k-1} or will be at the future t_{k+1}. The matrix elements defined by a set of successive measurements fit therefore "a posteriori" the evolution of the observable *V*, *i.e.* simply reproducing mathematically what is experimentally known in the considered time range. This empirical procedure, in principle non-predictive, is to be repeated at all times and extended to each observable of any event *K* of interest to characterize the whole system.

Moreover $\boldsymbol{K}$ implies neither past nor future: exchanging two lines, the change of sign of $\boldsymbol{K}$ is canceled by that of $\boldsymbol{A}$ concurrently necessary, *i.e.* $\boldsymbol{V}$ remains unchanged.

On the one hand this procedure, seemingly sterile, deserves attention as it shows that the link between numerical representation of the reality and physical events is in fact plausible: mathematics has its own rules to elaborate numbers; if these rules are implemented to reproduce the results of measurements, then the efforts of scientists are addressed to convert this empirical analysis of data, correct by definition, into rational information to be understood. So Wigner's doubts are bypassed regarding in fact the empiricism as an intermediate step between mere observation and profound knowledge of the reality, which however remains implicitly hidden in the raw data.

On the other hand all previous considerations evidence three key requirements necessary for any theoretical attempt to bridge abstract numbers and informative interpretation of results: 1) it must be holistic, 2) it must have space time structure, 3) it must inherently have evolutionary character. These three points prospect the non-trivial heuristic worth of $\boldsymbol{K}$: despite its pragmatic character, the coefficients of each line of the matrix and thus the matrix itself fulfill by definition these requirements and have thus physical valence. Also, $\boldsymbol{K}$ demonstrates the inherent rationality of Nature, without which no best fit technique could provide sensible outcomes. By consequence, no conceptual doubt exists about the effectiveness of a rational approach in describing mathematically the reality.

In principle is difficult to discern, on the basis of a linear combination of parameters only, whether for example two arbitrary time ranges δt and $\delta t'$ differ because they refer to different reference systems in reciprocal motion or because of the presence of a gravity field or even because the quantum uncertainty implies corresponding energy ranges $\delta\varepsilon$ and $\delta\varepsilon'$. Is evident thud the necessity of overcoming the mere empiricism hitherto preliminarily proposed, while acknowledging that the predictive ability of any theory is nothing else but its ability to reproduce the values of the aforesaid coefficients via rational path as general as possible, *i.e.* starting from first principles. In particular, it appears also necessary to identify rationally one by one the parameters f_j in fact concurring to describe exhaustively any physical event *K*.

The idea is at this point to bypass the best fit approach, valid by definition, by introducing a general function

$$\psi=\psi\left(x_{r,j},p_{r,j},\varepsilon_j,\cdots,\Phi_{r,j},t\right),\quad p_{r,j}=p_{r,j}\left(x_{r,j},t\right),$$
$$\varepsilon_j=\varepsilon_j\left(x_{r,j},t\right),\quad \Phi_{r,j}=\Phi_{r,j}\left(x_{r,j},t\right) \tag{1.3}$$

the index *r* stands for the set of three space coordinates and related vector components of all dynamical variables characterizing the system, e.g. possible internal and external vector fields $\Phi_{r,j}$ suitable to affect the evolution of all its constituting particles, the dots indicate any further *j*-th descriptive parameter additional to $p_{r,j}$ and ε_j possibly necessary. The last three positions allow writing implicitly and simply $\psi=\psi\left(x_{r,j},t_k\right)$ via the various $x_{r,j}$ of all *j*-th descriptive parameters contributing to the *k*-th line of ***K*** at the time t_k. For example $\varepsilon_j\left(x_{r,j},t\right)$ is itself a shortcut of $\varepsilon_j=\varepsilon_j\left(\Phi_{r,j},x_{r,j},t_k\right)$; indeed x_j at various t_k are somehow determined themselves by the strengths of the fields possibly acting on the system. So the Equation (1.3) can be shortened without loss of generality writing ψ as

$$\psi=\psi\left(X_{r,j},t_k\right),\quad X_{r,j}=x_j,y_j,z_j, \tag{1.4}$$

having nested into *X* all possible descriptive parameters implicitly governing the physical state of the system.

It is clear that the strategy of implementing the form (1.4) as a starting point, requires to extract successively from $X_{r,j}$ information about the possible external fields concurring to the internal interactions in defining $\Phi_{r,j}$ previously quoted. But how could the primordial function ψ summarize the variety of phenomena symbolized by every possible observable V_k for all possible physical events *K*?

Try to simplify the problem: although in principle the following considerations hold even for $r>3$, as postulated in some physical theories [3], assume for simplicity and without conceptual limitation a two dimensional space time, with the time coordinate and one space coordinate only. In this assumed one dimensional space $r=1$ can be omitted, whereas the space coordinates and respective vector components are represented by the unique index *j* that now refers to the various particles of the system. Accordingly, it is eventually possible to write more shortly $\psi=\psi\left(x_j,t_k\right)$ intending now *j* extended to the freedom degrees of all particles of a given physical system at the time t_k. So any physical effect determining the behavior of the system is described via one dimensional approach with two space time coordinates only for each freedom degree; this bypasses the difficulty of guessing one by one the descriptive parameters that effectively govern case by case the event *K*. Compare now the early empirical expression (1.1) with the series expansion of ψ around arbitrary initial coordinates x_{oj} and t_o, which reads

$$V_k=V_o+\sum_{j=1}^{J}\sum_{i=1}^{I}\frac{1}{i!}\left(\left(x_j-x_{oj}\right)\frac{\partial}{\partial x_j}+\left(t_k-t_o\right)\frac{\partial}{\partial t}\right)^i\psi\Bigg|_{x_j=x_{oj},t_k=t_o},\quad \psi=\psi\left(x_j,t\right) \tag{1.5}$$

the summation over *i* accounts for the arbitrary number *I* of terms of the series, that on *j* reproduces the same number of terms of the linear combination (1.1), the index k still represents the time at which the descriptive parameters *j* are expressed when defining the time change of an appropriate function ψ of all the necessary parameters. The notation indicates that the derivatives of ψ are calculated at arbitrary $x_j = x_{oj}$ and $t = t_o$ defining V_o in a given *R*, e.g. the laboratory. So each term *j* of (1.1) takes the form

$$A_j f_{t_k,j} = A_j f_{t_o,j} + \sum_{j=1}^{J}\sum_{i=1}^{I}\frac{1}{i!}\sum_{i'=0}^{i}\binom{i}{i'}\left(x_j - x_{oj}\right)^{i'}\left(t_k - t_o\right)^{i-i'}\partial_{oj}^{i'}\psi_o$$

$$\partial_{oj}^{i'}\psi_o = \frac{\partial^{i'}}{\partial x_j^{i'}}\frac{\partial^{i-i'}}{\partial t_k^{i-i'}}\psi\bigg|_{x_j = x_{oj},\, t_k = t_o} \tag{1.6}$$

the additive term is assumed known, being the initial boundary condition of the problem. Each *j*-th term is still related to the respective parameter f_j of the best fit procedure at the time t_k, although with a small difference. Previously $f_{t_k,j}$ were selected quantities implied by the physical event K (all measurable dynamical variables, among which x_j and t_k) tentatively introduced one by one; the best fit procedure aimed to calculate the respective coefficients A_j reproducing the known values of V_k (the specific physical property of interest) at the time t_k. Here instead the series expansion yields numerical coefficients $A_j = \partial_{oj}^{i'}\psi_o$ given by derivatives of a unique unknown function ψ calculated once for all at prefixed space and time coordinates initially set. The descriptive parameters are $(i,i')\, f_{t_k,j} = \delta x_j^{i'}\delta t_k^{i-i'}$, *i.e.* combinatorial factors times various products of space time ranges $\delta x_j = x_j - x_{oj}$ and $\delta t_k = t_k - t_o$: the dynamical variables previously tentatively introduced via the respective descriptive parameters $f_{t_k,j}$ correspond now to the space coordinates of all particles that still represent space and time experimental inputs. If these latter are known, then (1.6) and (1.1) are equivalent as concerns the best fit approach, defined again by a linear system of equations with best fit unknowns $\partial_{oj}^{i}\psi_o$. Yet, as by definition the coordinates depend upon all fields possibly acting on the system, summarized by $\Phi_{r,j}$ at given t_k and nested like in (1.4) and (1.3), further calculations are necessary to go back from these coordinates experimentally measured to the strength of the fields hidden in ψ. Nonetheless there is more information in (1.6) than in (1.1): the correlation of the actual experimental data to the initial conditions is not simply reasonable, it is required by the concept of space time ranges themselves.

The next step to overcome the legitimate Wigner doubts is just the time correlation (1.5), which does exist indeed and involves space time ranges as they appear in (1.6), not the local x_j, t_k and x_{oj}, t_o.

The worth of this information appears just from these equations comparing the particular cases where $i = 1$ and $i = 2$ in (1.6). Since in the former case $i' = 0,1$, the summation on i' yields for each *j* term $\partial_x \delta x_j + \partial_t \delta t_k$, where ∂_x and ∂_t are mere numerical coefficients corresponding to the respective $\partial_{oj}^{i'}\psi_o$. Whatever the numerical values of these coefficients might be, the space and time ranges appear separately: all δx_j on the one side and δt_k on the other side can be put independently equal to zero to describe local or simultaneous events. The case $i=2$ is conceptually different and more interesting, as the *j*-th term of (1.6) reads $\partial'_x \delta x_j^2 + \partial'_t \delta t_k^2 + 2\partial'_x \partial'_t \delta x_j \delta t_k$, being ∂'_x and ∂'_t new numerical coefficients; the space and time ranges appear together in the mixed term $\delta x_j \delta t_k$ with mixed coefficient $\partial'_x \partial'_t$. In general all higher order terms of the sum over $i>1$ imply mixed space time ranges.

Hence the first order and second order terms of the series do not imply merely two different degrees of numerical approximations in calculating V_k of the Equation (1.5). It is clear that $i=1$ is the classical case: a glance to this equation indicates that space and time terms are in fact separate dynamical variables like in (1.1). In the linear combination (1.1) the time is an independent input parameter, arbitrarily set, as a function of which the *x*-coordinate is next calculable consistently with any event occurring in the system, e.g. the interaction between particles. But in general the mixed terms modify strongly this point of view; for example it is no longer possible to put $\delta t_k = 0$ independently of δx_j: simultaneity and locality are in general conflicting concepts.

Moreover the Equation (1.6) introduces contextually the concept of evolution regarding in the same way also the initial configuration of the system through products of ranges $\delta x_{oj} \delta t_o$. So (1.6) shows that the local space and time coordinates separately measured and purposely introduced to carry out best fit calculations are actually mere mathematical parameters useful for empirical calculations only; the space time ranges of coordinates are instead physical parameters collecting together sets of local space coordinates x_j included within δx_j that define the evolution of allowed states of physical systems during a finite time lapse δt_k. Without this correlation, the system would be that of the matrix (1.2), *i.e.* describable as if it would consist of a list of mathematical terms unrelated and disconnected each other at various times.

This is the first hint to reproduce the coefficients A_j of (1.1) from first principles, thus overcoming both empiricism and Wigner's doubts.

In effect it will be found in the following that $\delta x_j \delta t_k$, not the local $x_j t_k$, is a sensible definition of space time compatible with quantum requirements. This shows that (1.6) lays prospectively the basis of both relativity and quantum physics: the necessity of a space time frame defined via sets of local coordinates $\delta x_j \delta t_k$ is in principle also consistent with the quantum lack of determinism based on local coordinates both exactly knowable.

Anyway, apart from mathematical details, the known value of any V_k in (1.2) is still reproducible in principle solving once more a set of linear equations of the unknown $\partial_{oj}^i \psi_o$. The expected rationality inherent the best fit calculation appears now through the mathematical properties of ψ. With a correct choice of this function, the coefficients $\partial_{oj}^{i'} \psi_o \leftrightarrow A_j$ describe conceptually and not only mathematically the evolution of physical systems; in practice this function still maps the systems like the mere empirical approach (1.2) and makes plausible the numerical representation of the reality. The key point is the underlying link with the concept of time evolution of physical systems with respect to their initial conditions.

Thus the basic idea is that a general function, ψ, must exist able to describe specific events of interest implementing the holistic concept of system evolution: if it is true that the Nature is a complex system under continuous modification, then the physical laws should also conform themselves to this principle. Accordingly, space and time should appear as inseparable properties in this evolutionary scenario that also implies the holistic view previously outlined as actual mathematical requirement.

On the one hand if the function ψ would be known, then there would be no necessity of determining in advance via best fit approach the power series (1.6) of the dynamical variables, which in fact would be calculable themselves through ψ and its derivatives; this chance exemplifies in principle the starting question of this section, *i.e.* to show why the rational knowledge of phenomena allows mapping the reality into numbers regardless of speculations about the geometrical origin of π. On the other hand this conclusion introduces the aim of the present paper, *i.e.* to understand how an appropriate function representing the physical phenomena through the concept of holistic evolution in fact prospects a conceptual path alternative to empirical best fit calculations; in this way ψ also removes the necessity of knowing in advance case by case the specific event to be described. Therefore the previous question about the mathematical structure of the reality overlaps to the following ones: "how all information codified in physical formulas is in fact deducible from ψ?" and also "are the current results of such theoretical basis susceptible of predictive outcomes prospecting the possible future Universe"?

Clearly the second question concerns the development of science and has heuristic valence in describing anything effectively allowed to happen in a changing Universe.

The purpose of the present paper is to highlight some straightforward hints towards this aim, *i.e.* how in principle could a single function ψ describe all variety of phenomena occurring in the Universe.

For simplicity and brevity of exposition the model is deliberately one dimensional: this choice does not represent a conceptual limit, it merely aims to simplify the theoretical approach with mathematical formalism as simple as possible. Also, the model purposely considers scalar quantities: for example v is the component of the velocity vector $\boldsymbol{v}$; analogous consideration holds for the component p of the momentum $\boldsymbol{p}$. These positions allow writing only $\psi = \psi(x,t)$ without subscripts. The time evolution of this function in a given *R* is therefore given by $\delta\psi = \psi(x+\delta x, t+\delta t) - \psi(x,t)$.

To add a further step forwards, consider more closely the particular space time interval introduced by (1.6)

$$\delta\ell_{st} = (x_j - x_{oj})(t_k - t_o) = \delta x_j \delta t_k \tag{1.7}$$

as $\delta\ell_{st}$ consists of two ranges, the first problem is how to define position and size of both δx_j and δt_k in an appropriate reference system *R*. For example the coordinates x_{oj} and t_o can be defined in order to fix the distances of one boundary of δx_j and δt_k on the respective axes, e.g. x_{oj} and t_o, from their common origin *O* of *R*, imagined as a two dimensional space time plane with the time on the vertical axis and the length on the horizontal axis; so x_j and t_k fix the sizes of the ranges. However a better chance exists in this respect: it is possible to introduce the following average values calculated via the boundary coordinates themselves of the ranges only

$$\overline{x_j}=\frac{x_{oj}+x_j}{2},\quad \overline{t_k}=\frac{t_o+t_k}{2}. \tag{1.8}$$

To describe self-consistently size and position of δx_j and δt_k in the space time plane, these mean values are defined on the respective axes of *R* as follows

$$v_c''\overline{t_k}=\delta x_j,\quad v_c'\delta t_k=\overline{x_j} \tag{1.9}$$

the first definition relates δx_j to the average time $\overline{t_k}$ needed for a hypothetical particle to travel through the whole range size, whatever it might be, the second definition relates δt_k to the displacement rate of its average coordinate $\overline{x_j}$ related to the position of both range boundaries only. Clearly these definitions need introducing two velocities v_c' and v_c'' compliant with the strategy of having defined mean values characteristic of both ranges only; if indeed just these definitions characterize size and position of both ranges in a self-consistent way, then any reference to *O*, and thus to *R*, is lost. In other words, replacing (1.9) into (1.7) neither x_{oj} nor t_o appear anymore explicitly in $\delta\ell_{st}=\left(v_c'/v_c''\right)\overline{t_k}\,\overline{x_j}$: these mean values of coordinates are in effect identically compatible with different x_{oj}' and t_o', *i.e.* with any other O'. Multiplying now side by side (1.9), one finds $v_c'v_c''\left(t_k^2-t_o^2\right)=x_j^2-x_{oj}^2$ *i.e.*

$$\delta\ell_{jk}^2=v_c^2t_k^2-x_j^2=v_c^2t_o^2-x_{oj}^2,\quad v_c^2=v_c'v_c''. \tag{1.10}$$

The actual value of v_c does not require in principle any specific hypothesis; is however interesting its particular value, necessarily constant without contradicting (1.10), consistent with $\delta\left(t_k^2\right)$ and $\delta\left(x_j^2\right)$ regardless of the reference system *R* where are defined δx_j and δt_k; the expression at the right hand side, formally identical to that at the left hand side, can be referred indeed to another reference system R_o. It is significant that a unique constant v_c^2 fits different time and space coordinates and that this equation implies different time and space ranges in different inertial reference systems *R* and R_o even in reciprocal motion.

However, the fact that space time terms $\delta x_j^i\delta t_k^{i'}$ more complex than that of (1.7) also appear in (1.6), suggests that a more complex space time metric is to be expected too. Since now all of these hints seem a reasonable step towards the special and general relativity. This also suggests that a model prospectively aimed to account someway for these suggestions should consider since the beginning not only δt and δx but also, at least, $\delta\left(\delta t\right)=\delta^2t=\delta t-\delta t'$ and $\delta\left(\delta x\right)=\delta^2x=\delta x-\delta x'$. Thus the problem is how to handle methodically both changes $\delta\psi$ and $\delta^2\psi$, rather than ψ itself, to describe systematically the physical properties of any system concerned by ψ. Despite ψ is not known, are essential and enough to this purpose the general definitions

$$\delta\psi = \psi(x+\delta x, t+\delta t) - \psi(x,t), \quad \psi = \psi(x,t) \tag{1.11}$$

and, increasing again $x+\delta x$ by δx and $t+\delta t$ by δt,

$$\delta(\delta\psi) = \delta^2\psi = \psi(x+2\delta x, t+2\delta t) - 2\psi(x+\delta x, t+\delta t) + \psi(x,t) \tag{1.12}$$

the former defines $\delta\psi/\delta\ell$, the latter $\delta^2\psi/\delta\ell^2$. Note that being by definition

$$\psi(\ell) = \psi(\ell_0) + \frac{\partial\psi}{\partial\ell}\delta\ell + \cdots, \quad \delta\ell = \ell - \ell_0,$$

where ℓ is any descriptive parameter of a physical system in the sense previously introduced, it is possible to put at the first order of approximation

$$\frac{\psi(\ell)-\psi(\ell_0)}{\ell-\ell_0} = \frac{\delta\psi}{\delta\ell} \approx \frac{\partial\psi}{\partial\ell} + \cdots$$

neglecting the higher order terms.

Here and in the following x and δx symbolize the *r*-th space coordinate of each *j*-th particle of the system and its change as a function of δt upon which depend possible changes of all dynamical variables and their *x*-components, e.g. $\delta\varepsilon_j$ and δp_j; the same holds for $\delta(\delta p_j)$ and $\delta(\delta\varepsilon_j)$, and so on. Although is considered for brevity and simplicity of notation one dimensional space coordinate only, from a conceptual point of view the number of actual coordinates is not necessarily limited to the usual three currently accepted.

The remainder of the paper concerns these points through an "*ab initio*" theoretical model whose exposition aims to be as self-contained as possible. Such model aims to deduce both well known results, as a validation, and new achievements, as innovative implications: in both cases, however, the assessment benchmark is its conceptual root in the Equations (1.11) and (1.12) only.

Despite for sake of brevity and clarity of exposition physical properties like energy and momentum have been taken for granted and explicitly mentioned as well acknowledged concepts in this introductory section, actually all of them will be inferred self-consistently themselves uniquely through (1.11) and (1.12); this holds also for quantities like charge and mass that apparently have nothing to do with the concept of evolution defined by these equations. Although seemingly trivial and innocuous, these two equations are unique source of information and unique input enough to infer all considerations exposed below in a consequential way, while overcoming Wigner's doubts and renouncing to any hints from physics theories currently existing. For completeness, when necessary, are also shortly sketched some results previously published to emphasize their connection with the present conceptual frame.

2. The Model

To infer information of physical interest from the initial positions (1.11) and (1.12), the simplest idea is to relate appropriately δx and δt, and possibly even $\delta^2 x$ and $\delta^2 t$, to $\delta\psi$ and $\delta^2\psi$ in an arbitrary *R*. In principle this correlation can be expressed implementing $\delta\psi$ to obtain two identities $\delta\psi = (\delta\psi/\delta x)\delta x = (\delta\psi/\delta t)\delta t$ that merge into

$$\frac{\delta x}{\delta t} = \frac{\delta\psi/\delta t}{\delta\psi/\delta x}. \tag{2.1}$$

The ratio at the left hand side introduces a new concept implied by $\delta\psi$, the velocity *v*; this dynamical variable, not evident nor necessary in (1.11) and (1.12), is defined by the identity

$$\frac{\delta\psi}{\delta x} = \frac{1}{v}\frac{\delta\psi}{\delta t}, \quad v = \frac{\delta x}{\delta t}. \tag{2.2}$$

The significance of this result, which follows the Equation (1.11) only, appears rewriting both sides according to the Equation (1.12) *i.e.* implementing likewise the identity $\delta^2\psi = \delta^2\psi$. Dividing both sides by δx^2 still via $\delta x = v\delta t$ just introduced, an analogous reasoning yields the further identity

$$\frac{\delta^2\psi}{\delta x^2} = \frac{1}{v^2}\frac{\delta^2\psi}{\delta t^2}. \tag{2.3}$$

The explicit physical meaning of these identities appears when $\delta \to \partial$, *i.e.* when the range sizes described by δ tend to zero. On the one hand this is possible because no restrictive hypothesis has been introduced about the ranges, on the other hand $\delta x \to 0$ and $\delta t \to 0$ do not necessarily imply equal limits $\partial\psi/\partial x$ and $\partial\psi/\partial t$ of the Equations (1.11). As written, the left hand side of (2.3) reads $\left[\psi(x+2\delta x,t) - 2\psi(x+\delta x,t) + \psi(x,t)\right]/\delta x^2$, whereas the right hand side reads $v^{-2}\left[\psi(x,t+2\delta t) - 2\psi(x,t+\delta t) + \psi(x,t)\right]/\delta t^2$. The limits of these expressions for $\delta \to \partial$ are indeed $\partial^2\psi/\partial x^2$ and $v^{-2}\partial^2\psi/\partial t^2$; as such, they are defined in general by the local analytical dependence of ψ upon either dynamical variable.

All this makes sense, as in fact the symbols δ indicate arbitrary changes not only of ψ but also of *x* and *t*; just for this reason, therefore, nothing can be "a priori" inferred from the ratios between $\delta\psi$ and δx or δt since both these latter are arbitrary, unknown, unrelated and thus implementable separately and independently each other. Instead, despite (2.3) is trivial identity,

$$\frac{\partial^2\theta}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\theta}{\partial t^2}, \quad \theta = \theta(x,t) \tag{2.4}$$

has physical meaning while the aforesaid limits imply contextually $\psi \to \theta$; the notation remarks that θ yields in particular the local analytical form of ψ resulting from the specific correlations of $\delta\psi$ with δx and δt (2.2) and (2.3). So the local behavior in the infinitesimal space range dx and time range dt fulfills at any x,t just the Equation (2.4). All quantities concerned by δ are arbitrary and finite by definition; thus they have been handled, and will be again handled also in the following, according to standard algebraic rules likewise any finite dynamical variable. Instead the limits $\delta \to \partial$ imposed to them define a further local condition/constrain that in fact eliminates their total arbitrariness and thus implies the mutual interdependence of both sides of (2.3) around a common limit: the initial analytic form of ψ, whatever it might be, turns locally into that, θ, fulfilling both local limits. In this specific case one has found the D'Alembert equation describing the dynamics of a homogeneous elastic string vibrating with fixed extremities and with constant propagation rate v of the perturbation around the equilibrium position of the string. Obviously the local dependence of θ upon x and t is found by solving the resulting differential equation.

This first example has emphasized how to infer information about one specific physical system through the local extrapolation θ of ψ as a function of both $\delta t \to 0$ and $\delta x \to 0$. Although the Equations (2.3) and (2.4) have identical analytical form, they remark the transition from non-local to local description of the concerned physical system: the former is in fact non-calculable, being mere identity, the latter takes physical meaning because is calculable and comparable with the experience. Otherwise stated, it is reductive to regard (2.3) as intermediate algebraic step towards (2.4); it actually describes a non-real and non-local world that does not have identifiable physical properties of the real and local world accessible to the experiment. Non-locality and non-reality are concurrent features of a further world, the quantum world, that can be not only guessed but also implemented to understand the microscopic properties of matter.

It is easy to generalize this result to the case where the string is non-homogeneous simply considering another possible chance of defining the link between $\delta\psi$ and both δx and δt via the trial positions $\delta x/k_1$ and $\delta t/k_2$, *i.e.* introducing two different proportionality factors k_1 and k_2 concerning separately the previous δx and δt. In this case $\delta\psi$ is defined via these generalized increments, both still unknown and arbitrary of course, where however the functions k_1 and k_2 prospect a new result even more general than (2.3). Now let us repeat the previous steps. To modify the correlation of $\delta\psi$ upon δx and δt via the respective factors k_1 and k_2, multiply first both sides of (2.2) by k_1, still keeping the definition $\delta x = v\delta t$ although with a different expectation value of the resulting local v. So the identity

$$k_1 \frac{\delta\psi}{\delta x} = \frac{k_1}{v}\frac{\delta\psi}{\delta t}, \quad k_1 = k_1(x,t) \tag{2.5}$$

yields the further identity according to (1.12)

$$\delta\left(k_1 \frac{\delta\psi}{\delta x}\right) = \delta\left(\frac{k_1}{v}\frac{\delta\psi}{\delta t}\right) \tag{2.6}$$

and thus, dividing both sides by δx,

$$\frac{\delta}{\delta x}\left(k_1 \frac{\delta\psi}{\delta x}\right) = \frac{\delta}{v\delta t}\left(\frac{k_1}{v}\frac{\delta\psi}{\delta t}\right), \quad \delta x = v\delta t \tag{2.7}$$

formally the Equation (2.7) results from two steps, taking first the changes (2.6) of the quantities at both sides of (2.5), which are subsequently related to δx and δt to obtain (2.7). Of course the limit $\delta \to \partial$ is not implemented at this intermediate step, as this would mean differentiating the quantities at both sides before having introduced the second function k_2; instead it is convenient to keep still finite changes of δx and δt, which again can be further worked out regarding them like any finite physical variable, to introduce k_2 too. Write thus without loss of generality

$$\frac{\delta}{\delta x}\left(k_1 \frac{\delta \psi}{\delta x}\right) = \frac{1}{v_0 v} \frac{\delta}{\delta t}\left(k_2 \frac{\delta \psi}{\delta t}\right), \quad \frac{k_1}{v} = \frac{k_2}{v_0}, \quad k_2 = k_2(x,t), \tag{2.8}$$

being v_0 a constant velocity by definition; therefore

$$\frac{\delta}{\delta x}\left(k_1 \frac{\delta \psi}{\delta x}\right) = \frac{1}{v'^2} \frac{\delta}{\delta t}\left(k_2 \frac{\delta \psi}{\delta t}\right), \quad v'^2 = v_0 v = v_0^2 \frac{k_1}{k_2}. \tag{2.9}$$

Now it is possible to infer from the Equation (2.9) the pertinent differential equation once more via the position $\delta \to \partial$ that implies thus a new local function $\psi \to \vartheta$, *i.e.*

$$\frac{\partial}{\partial x}\left(k_1 \frac{\partial \vartheta}{\partial x}\right) = \frac{1}{v'^2} \frac{\partial}{\partial t}\left(k_2 \frac{\partial \vartheta}{\partial t}\right), \quad \vartheta = \vartheta(x,t), \quad v' = v'(x,t). \tag{2.10}$$

The particular result with $k_2 = const$, which thus can be included in $k_1(x,t)$ at left hand side, yields the well known equation of the wave propagating through a non-homogeneous string with one fixed extremity. Obviously the functions θ and ϑ fulfilling the respective local limits implied by (2.4) and (2.10) are different; is indeed different the local behavior of either function correspondingly to the respective differential equations. The notation emphasizes that $\theta \neq \vartheta$: these functions describe different physical systems because of the different correlation of $\delta\psi$ with δx and δt.

The outcomes (2.4) and (2.10) highlight the strategy of the present paper: the arbitrary function ψ initially introduced according to (1.4) to describe in principle the physical properties of any system is implementable in various ways, depending on how is expressed the possible correlation between its change $\delta\psi$ with respect to that of its dynamical variables δx and δt. In other words the crucial point is not the analytical form of ψ, but how it changes as a function of δx and δt: whatever ψ might be, in fact this procedure identifies itself the possible kind of problem and outlines its mathematical solution as well via the resulting differential equation.

These results are not accidental outcomes inherent the explanatory examples just carried out; in effect no "ad hoc" hypotheses have been made on the concerned systems, e.g. homogeneous or non-homogeneous string, having simply introduced two different ways of describing the local change, *i.e.* the evolution, of ψ.

Let us exemplify further possible ways to handle $\delta\psi$ and $\delta^2\psi$ to confirm further the general worth of this strategy. To this purpose multiply both sides of (2.9) by v'^2 so that

$$v_0 v\frac{\delta}{\delta x}\left(\frac{k_1\delta\psi}{\delta x}\right)=\frac{\delta}{\delta t}\left(\frac{k_2\delta\psi}{\delta t}\right), \tag{2.11}$$

which suggests the following definitions according to (2.8)

$$v_0 v\frac{\delta p}{\delta x}=\frac{\delta\left(pv_0\right)}{\delta t}=\frac{\delta\varepsilon}{\delta t}, \quad p=k_1\frac{\delta\psi}{\delta x}, \quad \varepsilon=k_2\frac{\delta\psi}{\delta t}. \tag{2.12}$$

As the unique Equation (2.9) cannot specify both k_1 and k_2, which are still undefined, nothing excludes in principle the chances $k_2 \neq k_1$ or $k_2 = k_1$. Yet, even so, it appears that the positions (2.12) are not merely formal. To understand the physical meaning of the "new" quantities p and ε, note that the first equation implies

$$\delta\left(pv_0\right)=\delta\varepsilon, \quad i.e. \quad pv_0=\varepsilon+const, \quad k_1\neq k_2. \tag{2.13}$$

By definition $\delta\varepsilon=\varepsilon_2-\varepsilon_1$ and $\delta\left(pv_0\right)=p_2v_0-p_1v=0$, whereas $\varepsilon_1\leq\varepsilon\leq\varepsilon_2$ and $p_1v_0\leq pv_0\leq p_2v_0$; of course all quantities labeled "1" and "2" are arbitrary. As it possible to multiply side by side these equations, write $pv_0\delta\left(pv_0\right)=\varepsilon\delta\varepsilon+const\delta\varepsilon$. Intuitively $\varepsilon\delta\varepsilon$ should read $\delta\left(\varepsilon^2\right)/2$ with notation that avoids confusion between $\delta\varepsilon^2=\left(\varepsilon_2-\varepsilon_1\right)^2$ and $\delta\left(\varepsilon^2\right)=\varepsilon_2^2-\varepsilon_1^2$: in effect if ε is specifically regarded as mean value within its own allowed range $\delta\varepsilon$ of variability, *i.e.* $\varepsilon=\left(\varepsilon_2+\varepsilon_1\right)/2$, one finds $\varepsilon\delta\varepsilon=\left(\varepsilon_2+\varepsilon_1\right)\left(\varepsilon_2-\varepsilon_1\right)/2=\delta\left(\varepsilon^2\right)/2$ whatever ε_2 and ε_1 might be. The same reasoning for $pv_0\delta\left(pv_0\right)$ yields $\delta\left(pv_0\right)^2/2$. The idea of local variables ε and pv_0 allows an interesting implication noting that in general $\delta\varepsilon\equiv\delta\left(\varepsilon\pm const\right)$; so merging (2.13) one finds

$$\frac{1}{2}\delta\left(pv_0\right)^2=\frac{1}{2}\delta\left(\varepsilon\right)^2+\delta\left(const\,\varepsilon\right) \tag{2.14}$$

i.e. $\left(pv_0\right)^2=\varepsilon^2+2const\,\varepsilon$. Also, since $\varepsilon^2+2const\,\varepsilon=\left(\varepsilon+const\right)^2-const^2$, then

$$\varepsilon'^2=\left(pv_0\right)^2+const^2, \quad \varepsilon'=\varepsilon+const. \tag{2.15}$$

To examine either chance, calculate with the help of (2.12), (2.2) and (2.9)

$$\varepsilon+pv=k_2\frac{\delta\psi}{\delta t}+k_1\frac{\delta\psi}{\delta x}\frac{\delta x}{\delta t}=\left(k_1+k_2\right)\frac{\delta\psi}{\delta t}=\left(1+\frac{k_1}{k_2}\right)\varepsilon=\left(1+\frac{v}{v_0}\right)\varepsilon. \tag{2.16}$$

Regarding separately the addends at the initial and final left and right hand sides, this chain of equations is consistent: ε and pv at the left hand side correspond respectively to ε and $\left(v/v_0\right)\varepsilon$; in effect $p=\varepsilon/v_0$ is nothing else but (2.13) with $const=0$. This justifies regarding p and ε of (2.16) as momentum and energy in agreement with (2.23). Yet another chance also consistent with $k_1\neq k_2$ is

$$k_1=-k_2, \tag{2.17}$$

i.e. $\varepsilon+pv=0$ so that $v=-v_0$: in effect v is actually velocity component defined by δx during the time range δt.

The well known Equation (2.15) will be inferred again later; these short notes aim to justify preliminarily the positions (2.12) according which, regarding from now on $v_0 = c$ with usual notation, p and ε are nothing else but momentum and energy of a relativistic free particle. Simply regarding p and ε as local random values in their allowed ranges δp and $\delta\varepsilon$, *i.e.* anticipating here the concept of quantum uncertainty, it also appears in (2.14) why v_0 must be upper bound: if not, then p necessarily finite in its finite range δp could be consistent with an infinite energy ε' allowed by diverging $\delta\varepsilon$ once multiplied by a value of $v_0 \to \infty$.

So the finite value of c follows as a corollary.

Also, it is not surprising that the energy is defined an arbitrary constant apart; it will be shown shortly, however, that the constant has in this context a peculiar physical meaning. If $k_1 = k_2$, then $v = v_0$ and thus $\varepsilon + pv = 2\varepsilon$ *i.e.* $\varepsilon = pc$. The implications of this chance will be examined in the following.

2.1. Diffusion Equations

With $v = v_0 = c$, according to (2.8), the Equation (2.11) reads

$$\frac{\delta}{\delta x}\left(\frac{c^2 k_2 \delta\psi}{\delta x}\right) = \frac{\delta\psi'}{\delta t}, \quad \psi' = \frac{k_2 \delta\psi}{\delta t}, \quad k_1(x,t) = k_2(x,t); \tag{2.18}$$

since it is certainly possible to introduce an arbitrary function g such that $g\delta\psi = \tau\delta\psi'$, being τ a time dimensional constant, this equation reads

$$\frac{\delta}{\delta x}\left(D\frac{\delta\psi'}{\delta x}\right) = \frac{\delta\psi'}{\delta t}, \quad D = v_0^2 \tau \frac{k_2}{g}, \quad \delta\psi' = \frac{g}{\tau}\delta\psi, \quad g = g(x,t). \tag{2.19}$$

Whatever the function k_2/g might be, D has physical dimensions of diffusion coefficient; in effect with the position $\delta \to \partial$, which implies the local behavior of ψ' described by $\psi' \to \beta$, the last equation reads

$$\frac{\partial}{\partial x}\left(D\frac{\partial\beta}{\partial x}\right) = \frac{\partial\beta}{\partial t}, \quad D = D(x,t), \quad \beta = \beta(x,t). \tag{2.20}$$

This is just the general form of diffusion equation in a homogeneous and isotropic medium in the absence of internal sources or sinks. But diffusion of what? Although β is by definition dimensionless function, two relevant examples are reported below. To this purpose are anticipated here for clarity the concepts of mass and energy kT; both concepts will be inferred later self consistently in the frame of the present theoretical model.

It is possible to multiply β at both sides by a constant mass per unit volume m_0/V_0; so the equation

$$\frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x}\right) = \frac{\partial C}{\partial t}, \quad C = C(x,t) = \frac{\beta m_0}{V_0}, \tag{2.21}$$

where C is an appropriate function describing the local value of mass density, concerns the matter transport function under non-equilibrium concentration gradient. It is known that other important phenomena fulfill (2.20); in fact the extension to these cases, e.g. the Fourier heat diffusion, is also possible in an analogous way. Implementing a different dimensional factor to the local function β, *i.e.* multiplying both sides by an appropriate constant energy ϵ_0, one finds the famous equation

$$\frac{\partial}{\partial x}\left(\mathrm{K}\frac{\partial\epsilon}{\partial x}\right)=\frac{\partial\epsilon}{\partial t},\quad \epsilon(x,t)=\beta\epsilon_0=\beta kT_0,\quad \mathrm{K}=\mathrm{K}(x,t),\quad T=\beta T_0=T(x,t) \tag{2.22}$$

where now with usual notation K replaces D to express the heat diffusion coefficient simply identifying $\epsilon(x,t)\equiv kT$.

Note that the present strategy to infer information about physical systems reveals unexpected links between seemingly different laws: it is significant the fact that elementary manipulations of the equation of vibrating string lead to the diffusion equations.

2.2. Energy and Momentum

The Equation (2.8) reads according to (2.9)

$$\frac{\delta p}{\delta x}=\frac{1}{c^2}\frac{\delta\varepsilon}{\delta t},\quad p=k_1\frac{\delta\psi}{\delta x},\quad \varepsilon=k_1\frac{\delta\psi}{\delta t},\quad k_1=k_2,\quad v=c \tag{2.23}$$

because of course the positions (2.12) still hold also in this particular case. This equation can be implemented in two ways.

The first way is

$$\frac{\delta p}{\delta x}\frac{\delta t}{\delta\varepsilon}=\frac{\delta p}{\delta\varepsilon}\frac{1}{v}=\frac{1}{c^2} \tag{2.24}$$

and thus the second equality yields

$$\delta p=\frac{v}{c^2}\delta\varepsilon. \tag{2.25}$$

Here v still appears because the ratio $\delta x/\delta t$ is explicitly present in (2.24). The ranges explicitly written as $\delta p=p_2-p_1$ and $\delta\varepsilon=\varepsilon_2-\varepsilon_1$ by definition, where of course the quantities labeled with subscripts 1 and 2 are arbitrary, yield

$$p_2-\frac{v}{c^2}\varepsilon_2=p_1-\frac{v}{c^2}\varepsilon_1;$$

this result reads therefore

$$p=\frac{v}{c^2}\varepsilon,\quad p_1\le p\le p_2,\quad \varepsilon_1\le\varepsilon\le\varepsilon_2, \tag{2.26}$$

where p and ε are random values by definition included within the respective ranges.

The second way is highlighted rewriting (2.25) as

$$\frac{\delta(pc)}{\delta x}=\frac{\delta(\varepsilon/c)}{\delta t},$$

which yields

$$\delta p_o\delta x=\delta\varepsilon_o\delta t=const=n\hbar,\quad p_o=\frac{\varepsilon}{c},\quad \varepsilon_o=pc; \tag{2.27}$$

the constant $n\hbar$, required to fulfill products of different variables, will be justified soon below; the notation emphasizes that p_o and ε_o are not constants. Thus, in agreement with (2.2), one also finds

$$\delta\varepsilon_o = v\delta p_o, \quad v = \frac{\delta x}{\delta t}. \tag{2.28}$$

A few remarks help to simplify the notations in the following:

- the subscripts of δp_o and $\delta\varepsilon_o$ will be omitted as both ranges are arbitrary, so they actually symbolize any sizes of the respective δp and $\delta\varepsilon$;
- the velocities v and c are profoundly different, as the former is defined as ratio of two range sizes whereas the latter is a universal constant of the Nature;
- the definitions of two "new" quantities, momentum p and energy ε, have been guessed in (2.23) by dimensional reasons according to the constant $\hbar$, once having defined dimensionless the coefficients k_1 and k_2.

The lack of specific assumptions on p and ε, e.g. about the sizes of their allowed ranges, implies their physical definition on mere dimensional basis. At the moment n has been formally introduced in (2.27) as mere proportionality factor of a constant, $\hbar$; dimensional reasons are enough to justify this position. In the following, see next Equations (3.1) and (3.2), it will be shown that n is actually an arbitrary integer, whereas the pertinent reasoning will also explain why the physical laws need quantization. The Equation (2.27) is particularly interesting as it correlates the products $\delta x\delta p$ and $\delta\varepsilon\delta t$ of four ranges of different dynamical variables, regardless of the necessity of the position $\delta \to \partial$ and regardless of the range sizes; despite all changes of dynamical variables are arbitrary, the fact of having introduced a relationship between δx and δt implies the general and non-local character of this connection.

Consider now the Equations (2.28) and (2.26): the former concerns ranges, the latter local values. Let us show that relevant physical information is obtainable merging these equations. Multiplying side by side

$$\delta\varepsilon = v\delta p, \quad \varepsilon = \frac{c^2}{v} p \tag{2.29}$$

one finds

$$\varepsilon\delta\varepsilon = c^2 p\delta p; \tag{2.30}$$

thus (2.30) is compatible with

$$\varepsilon^2 + const' = c^2\left(p^2 + const''\right). \tag{2.31}$$

So follow three relevant equations

$$p = \varepsilon\frac{v}{c^2}, \quad \varepsilon^2 = (pc)^2 + const^2, \quad const^2 = const'' v_0^2 - const' \tag{2.32}$$

Introduce now the boundary condition $p = 0$ to which corresponds $\varepsilon = \varepsilon_0 \neq 0$, because in general the third equation is different from zero; strictly speaking, in effect, there is no reason to expect that ε_0 is necessarily null too. So this boundary condition yields $\varepsilon_0 = \pm const$; moreover it implies defining a "new" quantity m not yet explicitly mentioned hitherto although implicitly inherent the physical dimensions of p and ε, *i.e.*

$$\lim_{v\to 0}\frac{p}{v}=\frac{\varepsilon_0}{c^2}=m \tag{2.33}$$

so m is the rest mass. Calling c the constant velocity v_0, with usual notation, the last result reads thus

$$p=\varepsilon\frac{v}{c^2},\quad \varepsilon^2=(pc)^2+\left(mc^2\right)^2. \tag{2.34}$$

Clearly the particular case $\varepsilon=pc$ corresponds to $v=c$, which however requires $m=0$ in the second (2.34). It is immediate to verify that the two Equations (2.34) are consistent for $m\neq 0$, as they imply the Lorentz factor $\sqrt{1-v^2/c^2}$, whereas it also follows

$$\varepsilon^2=\frac{\left(mc^2\right)^2}{1-(v/c)^2},\quad p^2=\frac{(mv)^2}{1-(v/c)^2}. \tag{2.35}$$

The second equation is compatible with $\pm p$; this is not surprising because actually the component of p along an arbitrary direction can have both signs. Much more interesting is the analogous conclusion for $\pm\varepsilon$, which implies states of negative and positive energy separated by a gap 2ε.

Note that in addition to the concepts of mass, momentum and energy, follow from (1.11) and (1.12) the constancy of light speed and Lorentz transformations of energy and momentum.

A problem however arises now about why the first (2.34) is consistent with $pc=\varepsilon$ for $v=c$ whereas both (2.35) and the second (2.34) itself do not. A rational answer to this question will be given in the next Section 4.3. Note at the moment that the factor c^2/v of (2.34) yields $c^2/v\nu=length$, being ν an arbitrary reciprocal time; so, calling "wavelength" the new length λ defined in this way and multiplying both sides of the first (2.34) by ν^{-1}, one finds $\varepsilon/\nu=const=p\lambda$. Thus

$$\frac{\varepsilon}{\nu}=const,\quad p=\frac{const}{\lambda},\quad \lambda=\frac{c^2}{v\nu}, \tag{2.36}$$

where obviously $const=h$; so $\nu=c/\lambda$ is defined even for $v=c$. These positions, here reasonably guessed, are easily verified starting again from (2.34) rewritten as $p=\varepsilon\delta x/c^2\delta t$. With the help of (2.27), trivial manipulation turn equivalently this result into both forms

$$\frac{(pc)\delta(pc)}{\varepsilon}=energy=\frac{\hbar}{\delta t},\quad \frac{(\varepsilon/c)\delta(\varepsilon/c)}{p}=momentum=\frac{\hbar}{\delta x}.$$

In both cases, dimensional considerations confirm the validity of the three positions (2.36), regarding in particular $\delta x\leftrightarrow\lambda$: *i.e.* the range size δx corresponds to one or more momentum wavelengths, the range size δt corresponds to one or more frequency quanta. This suggests that actually $\lambda=n\lambda_o$, with $n=integer$, which formally is compatible with the constant appearing in (2.36) as $const=n\hbar$ as in effect it has been guessed in (2.27). Therefore it is possible to write, in agreement with (2.2),

$$\varepsilon_\omega = \frac{\hbar}{\delta t}, \quad \frac{1}{\delta t} = 2\pi\nu, \quad p_\omega = \frac{\hbar}{\delta x}, \quad \frac{1}{\delta x} = \frac{2\pi}{\lambda}. \tag{2.37}$$

It is interesting the fact that the Equations (2.36), pillars of quantum mechanics, are obtained contextually to the relativistic expressions of momentum, energy and rest mass.

2.3. Lagrange and Hamilton Equations

Write (2.28) as

$$\frac{\delta p}{\delta t} = \frac{\delta \varepsilon}{\delta x};$$

rewriting left hand side via (2.27) with $n = 1$ for simplicity, this equation reads then according to (2.23)

$$\frac{\delta}{\delta t}\left(\frac{\hbar}{\delta x}\right) = \frac{\delta}{\delta t}\left(\frac{\delta \varepsilon}{\delta x/\delta t}\right) = \frac{\delta \varepsilon}{\delta x}$$

and thus

$$\frac{\delta}{\delta t}\left(\frac{\delta \varepsilon}{\delta \dot{x}}\right) = \frac{\delta \varepsilon}{\delta x}, \quad \delta \dot{x} = \frac{\delta x}{\delta t}.$$

Also now the general concept of energy takes physical meaning via the limit $\delta \to \partial$, which implies $\varepsilon \to \phi$ as well; hence the result is

$$\frac{\partial}{\partial t}\left(\frac{\partial \phi}{\partial \dot{x}}\right) = \frac{\partial \phi}{\partial x}, \quad \phi = \phi(x, \dot{x}). \tag{2.38}$$

So ϕ is the particular local energy resulting from ε whose local behavior is described just by this equation. It is easy to realize that the resulting ϕ turns out to be Lagrangian energy. The most intuitive interpretation of ϕ compatible with both sides of (2.38) is indeed

$$\frac{\partial \phi}{\partial \dot{x}} = p, \quad \dot{p} = \frac{\partial \phi}{\partial x} = F, \quad \dot{p} = \frac{\partial p}{\partial t}; \tag{2.39}$$

as all of this is coherent with ϕ equal to energy, these equations define ϕ reasonably consistent with the Lagrangian $\phi = \mathrm{T} - \mathrm{U}$ of a physical system. In effect defining

$$\mathrm{S} = \int \phi \mathrm{d}t, \quad F = -\frac{\partial \mathrm{U}}{\partial x},$$

one finds

$$\frac{\partial \mathrm{S}}{\partial t} = \phi, \quad \frac{\partial \mathrm{S}}{\partial x} = \int \frac{\partial \phi}{\partial x} \mathrm{d}t = \int \frac{\partial p}{\partial t} \mathrm{d}t = p \tag{2.40}$$

in agreement with the well known definition of action S. Moreover, the second (2.39) yields

$$\dot{p} = \frac{\partial \phi}{\partial x} = -\frac{\partial \mathrm{U}}{\partial x} = \frac{\partial (\mathrm{T} - \mathrm{U})}{\partial x}, \quad \dot{x}\frac{\partial \phi}{\partial \dot{x}} = \dot{x}p = 2\mathrm{T} \tag{2.41}$$

owing to Euler's theorem of homogeneous functions. Hence

$$\dot{x}\frac{\partial\phi}{\partial\dot{x}}-\phi=\mathrm{H}=\mathrm{T}+\mathrm{U}. \tag{2.42}$$

It is immediate to conclude that (2.42) yields the Hamilton function.

As the Equations (2.3) and (2.9) have sensible implications, (2.4) and (2.10), whereas (2.39) and (2.42) allow describing correctly the dynamics of any particle, the present approach appears significant: a relationship between space and time ranges δx and δt has been established even without knowing anything about the initial ψ, simply admitting possible relationships between arbitrary $\delta(\delta x)$ and $\delta(\delta t)$. Once more, however, it is worth emphasizing that everything follows via (2.27) from (1.11) and (1.12) only.

Instead of attempting to explain some particular physical event on the basis of the intuition about its presumed theoretical foundation, we started from arbitrary changes of an introductory function, ψ, which is not "a priori" specified but rather is "a posteriori" identified case by case depending on its possible local change described by the analytical form of the pertinent differential equation.

2.4. The Group Velocity

In (2.2) v is defined by the time range δt necessary for a particle to travel ideally the range size δx. Note now that (2.1) reads formally

$$\frac{\delta x}{\delta t}=\frac{\delta\omega}{\delta\mathrm{k}},\quad \delta\omega=\frac{\delta\psi}{\delta t},\quad \delta\mathrm{k}=\frac{\delta\psi}{\delta x},$$

where ω and k are two "new" quantities called frequency and wave vector respectively; in this case the concept of velocity at the left hand side is different from that of (2.2). These definitions introduce a further concept of velocity, because at the local limit $\delta\to\partial$ one finds

$$\frac{\partial x}{\partial t}=\frac{\partial\omega}{\partial\mathrm{k}}. \tag{2.43}$$

It is immediate to show that also the positions (2.36), in particular the third one, allow calculating consistently the group velocity of a matter wave packet through the following simple chain of equations. Implementing $\delta\varepsilon$ and δp with the help of (2.28) one finds

$$v=\frac{\delta\varepsilon}{\delta p}=\frac{\delta\nu}{\delta\lambda^{-1}}=\frac{\delta 2\pi\nu}{\delta 2\pi\lambda^{-1}}=\frac{\delta\Omega}{\delta\kappa},\quad \kappa=2\pi\lambda^{-1},\quad \Omega=2\pi\nu \tag{2.44}$$

whatever $const$ might be. This suggest a possible quantum definition of velocity additional to the direct ratio between space range δx and time range δt. Once more the position $\delta\to\partial$ implies the local definitions $\Omega\to\omega$ and $\kappa\to\mathrm{k}$, whereas v turns to the first equality into local group velocity v_g, *i.e.*

$$v_g=\frac{\partial\omega}{\partial\mathrm{k}}. \tag{2.45}$$

Eventually, note that the third Equation (2.36) alone is enough itself to confirm this result. Write

$$\lambda=\frac{\mathrm{n}c}{v},\quad \mathrm{n}c=\lambda v=\frac{\omega}{\mathrm{k}},\quad \mathrm{n}=\frac{c}{v};\tag{2.46}$$

as $c\delta\mathrm{n}=\lambda\delta v+v\delta\lambda$, trivial manipulations yield

$$c\lambda\frac{\delta\mathrm{n}}{\delta\lambda}=\lambda^2\frac{\delta v}{\delta\lambda}+\mathrm{n}c=\mathrm{n}c-\frac{\delta v}{\delta\lambda^{-1}},$$

whence

$$\frac{\delta v}{\delta\lambda^{-1}}=\frac{c^2}{v}=\mathrm{n}c-c\lambda\frac{\delta\mathrm{n}}{\delta\lambda}=c\left(\mathrm{n}-\lambda\frac{\delta\mathrm{n}}{\delta\lambda}\right).$$

Therefore

$$v=\frac{c}{\mathrm{n}-\lambda\dfrac{\delta\mathrm{n}}{\delta\lambda}}$$

yields for $\delta\to\partial$ the local dispersion equation

$$v_g=\frac{c}{\mathrm{n}-\lambda\dfrac{\partial\mathrm{n}}{\partial\lambda}},\tag{2.47}$$

i.e. the well known group velocity of a matter packet wave.

In summary, relevant equations of physics are simply inferred and described through various chances of changing an arbitrary function ψ of time and space, regardless of its early specific physical meaning and without need of introducing initial hypotheses. This concerns crucially the functions k_1 and k_2 introduced in general in (2.8) and (2.9), whose specific analytical form determines the correlation of $\delta\psi$ with δx and δt: as it has been just highlighted, if $k_1=k_2$ in (2.9) then one obtains the diffusion Equation (2.10), if instead $k_1\neq k_2$ then one obtains further results concerned later thanks to the additional freedom degree allowed to k_2.

2.5. The Relativistic Velocity

The starting point is the first Equation (2.34), which must be rearranged in order to find a sum rule between two arbitrary velocities v_1+v_2 and their corresponding v_1' and v_2', e.g. in two different inertial reference systems R and R'. As $v^{-1}=\varepsilon/pc^2$, calculate first

$$\frac{\varepsilon_1}{p_1c^2}+\frac{\varepsilon_2}{p_2c^2}=\frac{1}{v_1}+\frac{1}{v_2}=\frac{1}{v_o},\quad v_o=\frac{v_1v_2}{v_1+v_2};$$

in this way one has introduced v_1+v_2 through the invariant momentum. It is necessary now to define in general v_o in a form suitable to relate v_1+v_2 and $v_1'+v_2'$, e.g. in another reference system. A reasonable position is the following linear combination that does not involve neither v_1v_2 nor $v_1'v_2'$, *i.e.*

$$v_o=\frac{c^2}{v_1'+v_2'}-\frac{c^2}{v_1+v_2},$$

so that the sought result is

$$v_1' + v_2' = \frac{v_1 + v_2}{v_1 v_2 / c^2 + 1}. \tag{2.48}$$

Accordingly any v summed to or subtracted from c still yields c.

3. Preliminary Implications of the Model

The results so far obtained are enough to get four relevant consequences, exposed below.

3.1. Statistical Formulation of Quantum Uncertainty

Write (2.27) as

$$\delta x \delta p = n \times const = \delta \varepsilon \delta t \tag{3.1}$$

being n an arbitrary integer. The reason of this definition is to make (2.27) independent of a specific reference system. Suppose that (3.1) holds for ranges defined in R whereas $\delta x' \delta p' = n' \times const$ holds for that defined in any R', with n' arbitrary integer as well; the prime symbols account for the respective Lorentz transformations of range sizes. Actually the reference systems are indistinguishable because neither n nor n' are specific numbers, they instead symbolize by definition whole sets of allowed integer numbers: so any specific n of the first set that turns into a new specific n' of the primed set does not imply in fact distinguishable sets of the respective reference systems. This point is better understood introducing appropriate measure units $x_{Pl}, p_{Pl}, \varepsilon_{Pl}, t_{Pl}$ to express the respective range sizes; for example it is possible to express the size of δx as $n_x^* \times x_{Pl}$; *i.e.* n_x^* is a dimensionless length expressing the actual range size in x_{Pl} units.

Is evident the hint to the well known Planck units, whose choice implies $x_{Pl} p_{Pl} = \hbar = \varepsilon_{Pl} t_{Pl}$ by definition. Without having introduced the gravity constant yet, this explicit reference appears here premature; it is enough to emphasize that the Planck units fulfill this equation by definition. The crucial fact is that introducing the dimensionless lengths $n_x^*, n_p^*, n_\varepsilon^*, n_t^*$, the couples x_{Pl}, p_{Pl} and ε_{Pl}, t_{Pl} fulfill the condition $x_{Pl} p_{Pl} = \varepsilon_{Pl} t_{Pl}$. In this way, dividing side by side with the Equations (3.1), one finds

$$\frac{\delta x}{x_{Pl}} \frac{\delta p}{p_{Pl}} = n = \frac{\delta \varepsilon}{\varepsilon_{Pl}} \frac{\delta t}{t_{Pl}}, \quad const = \hbar = x_{Pl} p_{Pl} = \varepsilon_{Pl} t_{Pl}.$$

It implies that with this choice of measure units, the statistical formulation of quantum uncertainty reads simply

$$n_x^* n_p^* = n = n_\varepsilon^* n_t^* \tag{3.2}$$

the stars indicate arbitrary real numbers, n is instead an arbitrary real integer number. This reasoning shows that in fact the Equations (2.27) hold regardless of any reference system; otherwise stated, the problem of specifying the reference system where are defined the four uncertainty ranges is physically meaningless, provided that the local dynamical variables are systematically replaced by respective uncertainty range totally unknown in any physical problem. This holds also for the derivatives, which are defined in the present model as mere ratios of uncertainty ranges arbitrary, unknown and conceptually unknowable: for example is meaningless to inquire whether δx refers to Cartesian or curvilinear or cylindrical reference frame. What is crucial in this reasoning is that the four starred numbers be not specifiable and unspecified in any physical problem formulated via the Equations (2.10); in short, the quantization of n is necessary to make (3.1) independent of any specific R. For clarity and self-contained exposition, this is shortly sketched in the next subsection. The results quoted here for completeness are reported more in detail elsewhere [4] [5].

3.2. The Old Quantum Mechanics

It is usually assumed that the quantum problems are tackled via the operator formalism of wave mechanics, introducing operators and wave equations. For comparison purposes, this section sketches very shortly results concerning one case where the wave equation can be exactly solved: the non relativistic hydrogenlike atom. The aim is to show that identical information is obtainable via a a "corpuscular approach", which does not require solving any wave equation; it is enough to replace $x \to \delta x$ and $p \to \delta p$, instead of $-\hbar i \partial\Psi/\partial x$, and proceed via elementary algebraic manipulations. These results help understanding how the relativity fits the conceptual frame so far outlined.

The starting point is the classical component of $\boldsymbol{M} = \boldsymbol{r} \times \boldsymbol{p}$ along an arbitrary direction defined by the unit vector $\boldsymbol{w}$ is $M_w = \boldsymbol{r} \times \boldsymbol{p} \cdot \boldsymbol{w}$. Consider thus

$$M_w = (\Delta\boldsymbol{r} \times \Delta\boldsymbol{p}) \cdot \boldsymbol{w} = (\boldsymbol{w} \times \Delta\boldsymbol{r}) \cdot \Delta\boldsymbol{p} = \Delta\boldsymbol{W} \cdot \Delta\boldsymbol{p}, \quad \Delta\boldsymbol{W} = \boldsymbol{w} \times \Delta\boldsymbol{r},$$

which introduces a range of possible values for M_w included in ΔM_w. If $\Delta\boldsymbol{p}$ and $\Delta\boldsymbol{W}$ are orthogonal, then $M_w = 0$; else, rewriting $\Delta\boldsymbol{W} \cdot \Delta\boldsymbol{p}$ as $(\Delta\boldsymbol{p} \cdot \Delta\boldsymbol{W}/\Delta W)\Delta W$ with $\Delta W = |\Delta\boldsymbol{W}|$, the component $\pm\Delta p_W = \Delta\boldsymbol{p} \cdot \Delta\boldsymbol{W}/\Delta W$ of $\Delta\boldsymbol{p}$ along $\Delta\boldsymbol{W}$ yields $M_w = \pm\Delta W \Delta p_W$. Thus, according to Equations (3.1), $M_w = \pm l\hbar$, being l the usual notation for the integer quantum number of angular momentum. So M_w is effectively a multi-valued quantized function because of the uncertainties initially postulated for $\boldsymbol{r}$ and $\boldsymbol{p}$. One component of $\boldsymbol{M}$ only is actually knowable; the same considerations for the y and x components would trivially mean changing $\boldsymbol{w}$.

Just this conclusion on the physical uniqueness of M_w suggests that the average values $\langle M_x^2 \rangle$, $\langle M_y^2 \rangle$ and $\langle M_z^2 \rangle$ should be equal; so the quantity of physical interest to describe the properties of quantum angular momentum is l, as a function of which M^2 is now inferred as well. The components averaged over the possible states summing $(l\hbar)^2$ from $-L$ to $+L$, where L is an arbitrary maximum value of l, yield $\langle M_i^2 \rangle = \sum_{l_i=-L}^{l_i=L} (\hbar l)^2 / (2L+1)$ *i.e.*

$$M^2 = \sum_{i=1}^{3} \langle M_i^2 \rangle = L(L+1)\hbar^2, \quad M_w = l\hbar. \tag{3.3}$$

Consider the quantum system formed by a particle in a central force field, e.g.

an electron around a nuclear charge; the concept of force will be justified in the conceptual frame of (1.12) and (1.11). Assuming the origin O of R on the nucleus, let $\varepsilon = p^2/2m - Ze^2/r$ be the classical electron energy, where m is the electron mass. As $p^2 = p_r^2 + M^2/r^2$, putting again $p_r \to \Delta p_r$ and $r \to \Delta r$, one finds

$$\varepsilon = \frac{\Delta p_r^2}{2m} + \frac{M^2}{2m\Delta r^2} - \frac{Ze^2}{\Delta r}. \tag{3.4}$$

Two numbers of states, *i.e.* two quantum numbers, are expected because of the radial and angular uncertainties. In effect the Equations (2.1) and the quantum M^2 yield $\varepsilon = n^2\hbar^2/2m\Delta r^2 + l(l+1)\hbar^2/2m\Delta r^2 - Ze^2/\Delta r$, which reads

$$\varepsilon = \varepsilon_o + \frac{l(l+1)\hbar^2}{2m\Delta r^2} - E_o, \quad E_o = \frac{Z^2e^4m}{2n^2\hbar^2}, \quad \varepsilon_o = \frac{\left(n\hbar/\Delta r - Ze^2m/n\hbar\right)^2}{2m}. \tag{3.5}$$

Minimize ε putting $\varepsilon_o = 0$, which yields

$$\Delta r = \frac{n^2\hbar^2}{Ze^2m} = \left(\frac{\alpha Z}{n}\right)^{-1} n\lambda_C, \quad \lambda_C = \frac{\hbar}{mc}, \quad \alpha = \frac{e^2}{\hbar c} \tag{3.6}$$

and thus $\varepsilon_{\min} = \left[l(l+1)/n^2 - 1\right]E_o/n^2$; so $l \le n-1$ in order to get $\varepsilon < 0$, *i.e.* a bound state. The reason of both ways to express Δr will be explained in the section 6. Here are of interest the electron energy levels and rotational energy of the atom as a whole around O

$$\varepsilon_{\min} = \varepsilon_{el} + \varepsilon_{rot}, \quad \varepsilon_{el} = -\frac{Z^2}{n^2}E_0 = -\frac{Ze^2}{2\Delta r}, \quad \varepsilon_{rot} = E_0\frac{Z^2l(l+1)}{n^4}E_0 = \frac{e^4m}{2\hbar^2}. \tag{3.7}$$

The physical meaning of Δr is related to the early Bohr radius, *i.e.* ε_{el} is due to charges of opposite sign delocalized within a diametric distance $2\Delta r$ apart. So n and l are properties of the phase space, *i.e.* numbers of allowed quantum states.

Consider now the identity $\Delta r/n\hbar \equiv \omega\Delta r/n\hbar\omega$. So it is consequently true that

$$\frac{2\pi\Delta r}{nh} = \frac{v}{\epsilon} = \frac{1}{p}, \quad v = \omega\Delta r, \quad \epsilon = n\hbar\omega,$$

where the last equation of the chain introduces the momentum p by dimensional reasons and reads

$$2\pi\Delta r = n\lambda, \quad p = \frac{h}{\lambda}. \tag{3.8}$$

It shows the link between De Broglie momentum, Planck energy and condition $n\lambda = 2\pi\Delta r$, according which an integer number of steady electron wavelengths λ is defined along a circumference of radius Δr along which the electron wave propagates at rate v. For such electron waves one finds

$$\varepsilon_{el} = -\alpha\frac{Z}{n}\frac{pc}{2} = -\left(\frac{\alpha Z}{n}\right)^2\frac{mc^2}{2}. \tag{3.9}$$

The first chain of equalities will be explained in the next section 6, in particular as concerns the evident link of pc and mc^2 with E_0. Note here that introducing α to express the quantum energy levels implies defining the De Broglie momentum as a corollary, in agreement with (2.36) and (2.37): appears interesting that the energy levels ε_{el} of the system are linked to the kinetic energy pc of the electron moving along the circumference of radius Δr via the coefficient $\alpha Z/2n$. On the one hand, this result emphasizes the electromagnetic character of the interaction between electron and nucleus; comprehensibly Δr is proportional to α^{-1}, as the coupling constant determines the force exerted in an interaction, *i.e.* the greater α the smaller Δr. On the other hand, it also appears that the key role of the quantum uncertainty in determining the allowed energy levels (3.7) also evidences the kind of interaction itself.

These results confirm that the operator formalism and the uncertainty equations are equivalent in describing the quantum systems. As concerns the spin, the paper [6] [7] has shown that it can be inferred without additional hypotheses from the quantization itself. Simply rewriting identically $M^2 = (l+1/2)^2\hbar^2 - (\hbar/2)^2$, one finds

$$M^2 + (\hbar/2)^2 + \left(l+\frac{1}{2}\right)\hbar^2 = M^2 = \left(l+\frac{1}{2}\right)\left[\left(l+\frac{1}{2}\right)+1\right]\hbar^2 \qquad (3.10)$$

after having added $(l+1/2)\hbar^2$ at both sides. Trivial manipulations of the initial $\boldsymbol{M}$ exposed in the quoted paper show that

$$\mathcal{M}^2 = J(J+1)\hbar^2, \quad J = l_{or} + l_s, \quad l_s = \frac{1}{2} \qquad (3.11)$$

and that in general these consideration introduce the spin component $l'/2\hbar$; being of course l' an arbitrary integer, the quantum uncertainty implies itself the existence of bosons and fermions. No information is necessary about Δr and Δp_r, which in effect are unknown and unknowable because of the quantum uncertainty.

Besides its inherent worth, the hydrogenlike model has been explicitly quoted here because it also provides useful information about the characteristic lengths in the atom, the first of which is of course the Bohr radius inferred in (3.7). The first powers of α scale further significant lengths starting from this radius, whose essential form reads $r_B = \hbar^2/e^2m$ as a function of the fundamental constants. One infers the following lengths

$$r_B = \frac{\hbar^2}{e^2m}, \quad r_B\alpha = \lambda_C = \frac{\hbar}{mc}, \quad r_B\alpha^2 = r_e = \frac{e^2}{mc^2}, \quad r_B\alpha^3 = r_N = \frac{e^8}{\hbar^3mc^5} \qquad (3.12)$$

whose values are

$$r_B \approx 5.3\times10^{-9}\ \text{cm}, \quad \lambda_C \approx 3.6\times10^{-11}\ \text{cm},$$
$$r_e \approx 2.8\times10^{-13}\ \text{cm}, \quad r_N \approx 2.4\times10^{-15}\ \text{cm}.$$

the Bohr radius scales r_B down to λ_C, electron Compton length, and then to r_e, classical electron radius. Further lengths, shorter and shorter, will be introduced later to extend these definitions and sketch short range nuclear forces. Indeed the fact of having found these well known specific lengths suggests that even the fourth position should reasonably have its own physical meaning at the smaller α^3 scale too; if so, r_N can be related to nothing else but the scale of lengths within the atomic nuclei, whose sizes in effect are known to fall between $\lesssim 10^{-13}$ cm (proton of hydrogen) to $\sim 10^{-12}$ cm (heavier nuclei).

3.3. Velocity Dependence of Mass

Owing to (2.28), $\varepsilon_2 - \varepsilon_1 = vp_2 - vp_1$ reads $\varepsilon_2 - vp_2 = const = \varepsilon_1 - vp_1$; so (2.35) yields

$$\varepsilon - vp = \frac{mc^2 - mv^2}{\sqrt{1 - v^2/c^2}} = const.$$

This result is more expressively rewritten in the form $mc^2\sqrt{1 - v^2/c^2} = const$ fulfilled by

$$m_0c^2 = const, \quad m_0 = m\sqrt{1 - v^2/c^2}, \quad m = m(v). \tag{3.13}$$

The physical meaning of this result, the dependence of *m* on *v* via the constant m_0, will be clarified soon below. In the following are introduced three interesting ways to implement further the Equations (2.25) and (2.34), to show in particular how results of special relativity are obtainable regarding the local dynamical variables ε and *p* as random and unknown values defined in the respective quantum uncertainty ranges $\delta\varepsilon$ and δp.

3.4. Quantum Correction to Special Relativity

The strategy is still that followed to find (2.15) and to infer (2.31) and (2.32) from (2.29). Consider the Equation (2.25) and (2.26) rewritten in the particular case $v = c$ as

$$\delta\varepsilon^* = c\delta p, \quad \varepsilon^* = pc, \quad \varepsilon_1^* \le \varepsilon^* \le \varepsilon_2^*, \quad p_1 \le p \le p_2; \tag{3.14}$$

the former equation defines the maximum energy range $\delta\varepsilon^*$ allowed to the local ε^* consistently with the given momentum range δp allowed to any local *p*. Here energy and momentum ranges are linked each other, whereas in fact they were independent in the Equation (2.28) owing to the arbitrariness of *v*; so, the upper limit allowed for *v* implies an upper limit to the size of $\delta\varepsilon^*$ compliant with any possible δp. Anyway this latter is arbitrary; thus both energy and momentum ranges are in fact arbitrary as well, but now correlated. The second position emphasizes the local dynamical variables ε^* and *p* allowed in the respective ranges. The fact that (3.14) is not mere formal way of rewriting (2.25) but contains additional physical information, is easily proven: multiplying side by side both (3.14) one finds $\delta\left(\varepsilon^*\right)^2 = c^2\delta\left(p\right)^2$ *i.e.*, as in (2.14), $\varepsilon^{*2} + const' = \left(pc\right)^2 + const''$. So the second (2.34) is instantly inferred via the correlation between δp and $\delta\varepsilon^*$ through *c*.

However just the fact that the (3.14) appears suitable to be directly linked to (2.34) rises a quantum problem. Replace (2.36) in the Equation (2.34) via the positions $\varepsilon^* = hv^*$ and thus $p = h/\lambda^*$, being $\lambda^* = c/v^*$ in fact implied itself by the third (2.36) for $v = c$ too. Then $\left(hv^*\right)^2 = \left(hc/\lambda^*\right)^2 + \left(mc^2\right)^2$ requires

$m=0$. On the one hand nothing hinders in principle to express (2.34) via the corresponding quantum energy and momentum, in agreement with the dual wave/corpuscle character of matter. On the other hand (2.34), as written, seems inadequate to allow both $m \neq 0$ and (2.36). It is reasonable to expect that further terms to be included in (2.34) could overcome this difficulty: the attempt to generalize the standard result of the early special relativity is not only legitimate but also necessary.

The subsection 3.2 has been explicitly enclosed in the present exposition to emphasize that the quantum eigenvalues leave out any information about the range sizes; the Equations (3.3) to (3.12) elucidate this assertion. In other words the previous results obtained implementing $\delta p = p_2 - p_1$ could have been identically obtained considering any other $\delta p' = p_2' - p_1'$, as the range boundary coordinates are inessential as concerns the quantized eigenvalues of angular momentum and energy. The same holds of course even implementing a linear combination of momentum ranges, e.g. $\delta p'' = p_2'' - p_1'' = a\delta p + b\delta p'$ via the constant arbitrary coefficients *a* and *b*. This means that the local value *p* defined by $p_1 \le p \le p_2$ could be identically replaced by any p' defined by $p_1' \le p' \le p_2'$; the same holds of course for any p'' defined by $ap_1 + bp_1' \le p'' \le ap_2 + bp_2'$: the only essential requirement is that any range sizes δx and δp fulfill (3.1), whatever the boundary values might be. Now let us introduce in the relativistic domain this peculiarity of the quantum world. This means that the local values of pc and ε^* defined the respective ranges (3.14) can be replaced by linear combinations of momentum and energy.

The chance of demonstrating the actual effectiveness of this reasoning has heuristic worth in demonstrating the close connection between quantum and relativistic theories.

In practice, to generalize the standard relativistic result (2.34), implement again the first (3.14) with the same steps from (2.29) to (2.31) and then to (2.32), but rewriting the third and fourth positions as

$$\varepsilon_1^* \le \varepsilon_0 + \sigma_\varepsilon \varepsilon^* + \sigma_{o\varepsilon} \varepsilon^{*2} \le \varepsilon_2^*, \quad p_1 \le p_0 + \sigma_p p + \sigma_{op} p^2 \le p_2$$

σ_ε and σ_p are dimensionless arbitrary constants, *a* and *b* are arbitrary constants having physical dimensions $mass^{-1}$ and expressing conveniently the σ_o coefficients. The equations to be implemented are thus

$$\delta\varepsilon^* = c\delta p, \quad \sigma_\varepsilon \varepsilon^* - \frac{b}{c^2}\varepsilon^{*2} = \sigma_p pc - ap^2 + \epsilon_a, \tag{3.15}$$

where in fact $a \neq b \neq 0$ extend the previous procedure simply introducing additional ε^{*2} and p^2 terms with respect to (3.14) while however keeping a physical meaning still compliant with that of ranges $\delta\varepsilon^*$ and δp, as it appears via dimensional considerations. In other words, the second (3.15) still has the usual form $\delta\mathcal{E}^* = \delta(\mathcal{P}c) + const$. Repeat therefore exactly the same procedure just outlined to merge (3.14), *i.e.* multiply side by side the second and first (3.15) with the *a* and *b* terms exchanged of place; omitting for simplicity of notation the asterisk, one finds

$$\sigma_\varepsilon \varepsilon \delta\varepsilon + ap^2 \delta\varepsilon = bp\varepsilon\delta p + \sigma_p pc^2 \delta p + \delta\left(\epsilon_a pc\right), \tag{3.16}$$

which yields

$$\frac{1}{2}\delta\left(\sigma_\varepsilon \varepsilon^2\right) = \frac{1}{2}\delta\left(\sigma_p \left(pc\right)^2\right) - ap^2\delta\varepsilon + bp\varepsilon\delta p + \delta\left(\epsilon_a pc\right)$$

and then

$$\frac{1}{2}\delta\left(\sigma_\varepsilon \varepsilon^2\right) = \frac{1}{2}\delta\left(\sigma_p \left(pc\right)^2\right) - \delta\left(ap^2\varepsilon\right) + \delta\left(\epsilon_a pc\right), \quad b = -2a. \qquad 3.17)$$

Hence, reasoning as before, this result implies:

$$\sigma_\varepsilon \varepsilon^2 + const' = \sigma_p \left(pc\right)^2 + const'' - 2ap^2\varepsilon + const_1 + \epsilon_a pc + const_2. \tag{3.18}$$

As hold for (3.18) the same considerations carried out for (2.34), because also the new terms $ap^2\varepsilon$ and $\epsilon_a pc$ vanish for $p \to 0$. Merging the constants, one finds

$$\begin{gathered} \sigma_\varepsilon \varepsilon^2 = \sigma_p \left(pc\right)^2 - 2ap^2\varepsilon + \epsilon_a pc + \left(mc^2\right)^2 + \epsilon_0^2 \\ const'' - const' = \left(mc^2\right)^2, \quad const_2 + const_1 = \epsilon_0^2. \end{gathered} \tag{3.19}$$

The notation $\left(mc^2\right)^2$ has been kept resulting from the primed constants like in (2.32), in order that this equation reduces to (2.34) in the particular case $a = 0$ and $\epsilon_a = 0$. Of course the constants σ_ε and σ_p can be included in the respective energies; *i.e.* with the positions

$$\begin{gathered} p' = \sqrt{\sigma_p}\, p, \quad \varepsilon' = \sqrt{\sigma_\varepsilon}\, \varepsilon, \quad a' = \frac{a}{\sigma_p \sqrt{\sigma_\varepsilon}}, \\ \epsilon_a' = \frac{\epsilon_a}{\sqrt{\sigma_p}}, \quad m'c^2 = \sqrt{\left(mc^2\right)^2 + \epsilon_0^2} \end{gathered}$$

(3,19) reads

$$\varepsilon'^2 = (p'c)^2 + \left(m'c^2\right)^2 - 2a'p'^2\varepsilon' + \epsilon_a' p'c. \tag{3.20}$$

As expected, thanks to the higher order terms ε^2 and p^2 in (3.16) one finds again an equation like the second (2.34) plus two additional terms $-2a'p'^2\varepsilon'$ and $\epsilon_a' p'c$ not present in the standard special relativity. The quantum correction terms are negligible in (3.20) if $-2a'p'^2\varepsilon' + \epsilon_a' p'c \ll \varepsilon'^2$, *i.e.* if

$$-2\frac{p'^2/m_a}{\varepsilon'} + \frac{\epsilon_a' p'c}{\varepsilon'^2} \ll 1, \quad m_a = \frac{1}{a'}, \tag{3.21}$$

then ϵ_a and m_a fix the scale where the quantum correction plays a significant role. Moreover, if in particular $\epsilon_a' \ll 2a'p'\varepsilon'/c$, then is effective only the term $-2a'p'^2\varepsilon'$ in (3.20). These points deserve attention.

First of all, replace $\varepsilon' = h\nu'$ and $p' = h/\lambda'$; being again $\left(h\nu'\right)^2 = \left(hc/\lambda'\right)^2$, (3.20) reads

$$\begin{gathered} \left(m'c^2\right)^2 = 2a'\frac{h^2}{\lambda'^2}h\nu' - \epsilon_a'\frac{hc}{\lambda'} = h\nu'\left(2a'\frac{h^2}{\lambda'^2} - \epsilon_a'\right) = h\nu'\left(2m_a c^2\frac{\lambda_a^2}{\lambda'^2} - \epsilon_a'\right) \\ \lambda_a = \frac{h}{m_a c}, \quad m_a \ge \frac{\epsilon_a}{2} \end{gathered}$$

hence (3.20) is compatible with the quantum condition (2.36) even for $m'\neq 0$, whereas $m'=0$ is also possible if in particular vanishes the quantity in parenthesis. The last inequality holds for $\lambda'=\lambda_a$, because in general $\lambda'\geq\lambda_a$, whatever λ' might be.

Moreover rewrite the second (3.15) with the help of (2.36) as

$$\sigma_\varepsilon h\nu+\frac{2}{m_a c^2}(h\nu)^2=\sigma_p h\nu-\frac{h^2}{m_a\lambda^2}+\epsilon_a,\quad \nu=\frac{c}{\lambda}. \tag{3.22}$$

To recognize the physical meaning of this equation under the condition that m_a is the constant mass defined by the Equation (3.21), useful positions are:

$$\frac{h^2}{m_a\lambda^2}=m_a c^2\frac{\lambda_a^2}{\lambda^2},\quad h\nu=rm_a c^2,\quad \lambda=n\lambda_a,\quad r=n'\sigma_\varepsilon. \tag{3.23}$$

The first one is an identity, whose left hand side is simply rewritten introducing the Compton length λ_a of m_a. The second one is a formal way to link $h\nu$ and $m_a c^2$ via the parameter r to be defined. The third one regards λ of (3.22) as an integer multiple of λ_a; in fact the conceptual difference between p defined by (2.23) and $p=h/\lambda$ is that δx is a mere space range that can take in principle any value, the wavelength λ requires introducing quantized lengths $n\lambda$, which explains why anyway the quantization must be introduced via h in $n\hbar/\delta x$ of (3.1). Although this idea is introduced here as a reasonable input, a previous paper [6] has shown that in effect a huge amount of interesting results is accordingly obtainable. The fourth one will be explained after having replaced the first three (3.23) into (3.22), which reads

$$rh\nu+\sigma_\varepsilon\frac{h\nu}{2}=\frac{m_a c^2}{2}\left(r\sigma_p-\frac{1}{n^2}\right)+\frac{\epsilon_a}{2}. \tag{3.24}$$

For sake of generality the notation emphasizes that n' defining r is not necessarily coincident with n defining the ratio λ/λ_a. Is attracting the fourth position (3.23) with n' arbitrary integer that expresses the left hand side as σ_ε times the harmonic oscillator energy; indeed (3.24) becomes

$$n'h\nu+\frac{h\nu}{2}=\frac{m_a c^2}{2}\left(n'\sigma_p-\frac{1}{\sigma_\varepsilon n^2}\right)+\frac{\epsilon_a}{2\sigma_\varepsilon}.$$

Now it is necessary to express the fact that m_a is a constant, which in fact means regarding the quantum numbers n and n' as proportionality factors linking $m_a c^2$ and $h\nu$. The limit $n'\to\infty$ yields $n'h\nu=(n'\sigma_p/2)m_a c^2$ and thus, by comparison with the second and fourth positions (3.23), $\sigma_p/2=\sigma_\varepsilon$; so the last equation reads

$$n'h\nu+\frac{h\nu}{2}=\frac{m_a c^2}{2}\left(2n'\sigma_\varepsilon-\frac{1}{\sigma_\varepsilon n^2}\right)+\frac{\epsilon_a}{2\sigma_\varepsilon},\quad \sigma_p=2\sigma_\varepsilon. \tag{3.25}$$

It appears that if $m_a=0$, then $\epsilon_a/2\sigma_\varepsilon$ is the energy of harmonic oscillator of frequency ν. Analogous conclusion holds if $\sigma_\varepsilon^2=1/2(n'n)^2$, in which case $\epsilon_a/2=n'h\nu_0+h\nu_0/2$ with $\nu_0=\nu/\sqrt{2}n'n$; as both n and n' are arbitrary integers, $n'n$ must be regarded as a new arbitrary integer itself and thus anyone among the numbers already implied by n^2 and n'^2. So ν is an arbitrary multiple of the fundamental frequency ν_0. The fourth position also allows expressing (3.24) as a function of quantum numbers only

$$n'^2\sigma_\varepsilon + \frac{n'\sigma_\varepsilon}{2} = \frac{1}{2}\left(2n'^2\sigma_\varepsilon - \frac{1}{\sigma_\varepsilon n^2}\right) + \frac{1}{\sigma_\varepsilon}\frac{\epsilon_a}{m_a c^2},$$

which yields

$$\frac{\epsilon_a}{m_a c^2} = \frac{1}{2}\left(n'\sigma_\varepsilon^2 + \frac{1}{n^2}\right). \tag{3.26}$$

Next, inserting the positions (3.23) in (3.24) trivial manipulations yield

$$\left(\frac{m'}{m_a}\right)^2 = n'\sigma_\varepsilon\left(\frac{2}{n^2} - \frac{\epsilon_a}{m_a c^2}\right) = n'\sigma_\varepsilon\left(\frac{3}{2n^2} - \frac{1}{2}n'\sigma_\varepsilon\right)$$

Clearly ν appearing in (3.24) implied by $a' \neq 0$ and $\epsilon_a \neq 0$ is different from ν^* previously found consistent with $m = 0$ only; (3.20) skips this restriction.

It is known that (3.20) is a valuable equation of quantum gravity able to solve three cosmological paradoxes [8]. It is hard to guess what has to do the cosmology in this conceptual frame; but in fact this is not the correct way to regard this equation. Rather it is correct to say that the additional terms due to $a \neq 0$ and $b \neq 0$ add a quantum correction to the standard relativistic formula, actually having quantum character itself being inferred from (2.30) and (2.31); then, once having acknowledged this result, further studies also acknowledge that this correction has valuable cosmological implications as well.

3.5. Operator Formalism

The subsection 3.2 has shown that the corpuscular approach to quantum mechanics provides sensible results in agreement with the wave formalism. This subsection shows that also the wave formalism enters in the conceptual frame hitherto exposed. Implement the quantum relativistic Equation (3.20), noting that

$$\varepsilon''^2 = \left(mc^2\right)^2 + \left(pc\right)^2 = \left(mc + ip\right)\left(mc - ip\right)c^2, \quad \varepsilon''^2 = \varepsilon^2 + 2ap^2\varepsilon - \varepsilon_a pc. \tag{3.27}$$

Admitting that even the single factors at the right hand side have physical meaning, it is possible to introduce imaginary momentum $\mathcal{P}$ and energy $\mathcal{E}$ in agreement with the early positions (2.9) and (2.12); the momentum and energy equations take indeed the forms

$$\mathcal{P} = \mp i\hbar\frac{\delta\psi}{\delta x}, \quad \mathcal{E} = \pm i\hbar\frac{\delta\psi}{\delta t}, \quad \mathcal{E}^2 = \varepsilon''^2, \tag{3.28}$$

being simply required

$$k_1 = \mp ik_0, \quad k_2 = \pm ik_0, \quad k_0 = \hbar.$$

The correct correspondence of signs in (3.28) is indeed such that $k_1 + k_2 = 0$

and thus $\mathcal{P}v+\mathcal{E}=0$, in agreement with (2.17). Whatever the specific form of ψ might be, replacing p and ε'' of (3.27) with the new definitions (3.28) one finds

$$\left(mc+\hbar\frac{\delta\psi}{\delta x}\right)\left(mc-\hbar\frac{\delta\psi}{\delta x}\right)=-\left(\frac{\hbar}{c}\frac{\delta\psi}{\delta t}\right)^2; \tag{3.29}$$

in this way the Equation (3.27) turns again into a real form. Introduce now the positions

$$\frac{\delta^2\psi}{\delta t^2}=\pm\left(\frac{\delta\psi}{\delta t}\right)^2, \quad \frac{\delta^2\psi}{\delta x'^2}=\pm\left(\frac{\delta\psi}{\delta x'}\right)^2, \tag{3.30}$$

being

$$\delta x'=c\delta t, \tag{3.31}$$

which are justified soon below. In principle the positions (3.30) are compatible each other because δt and $\delta x'$ are arbitrary finite ranges that can be determined in order to fulfill both equations. Note that the more general positions $\delta^2\psi/\delta t^2=q\left(\delta\psi/\delta t\right)^2$ and $\delta^2\psi/\delta x^2=q\left(\delta\psi/\delta x\right)^2$, with q arbitrary factor, would have been in principle reasonable and possible; however q could be included in m of (3.29), so its specific value is inessential; more important are instead the signs of q, as it will appear shortly. Taking the upper signs (3.30), (3.29) reads as follows

$$\left(\frac{mc}{\hbar}\right)^2=\frac{\delta^2\psi}{\delta x^2}-\frac{1}{c^2}\frac{\delta^2\psi}{\delta t^2}-\frac{1}{v^2}\frac{\delta^2\psi}{\delta t^2}+\frac{1}{v^2}\frac{\delta^2\psi}{\delta t^2}.$$

The addend $\delta^2\psi/v^2\delta t^2$ has been summed and subtracted at the right hand side in order to split this equation as follows

$$\frac{\delta^2\psi}{\delta x^2}-\frac{1}{v^2}\frac{\delta^2\psi}{\delta t^2}=0, \quad \left(\frac{mc}{\hbar}\right)^2=-\frac{1}{c^2}\frac{\delta^2\psi}{\delta t^2}+\frac{\delta^2\psi}{\delta x^2}, \quad \delta x^2=v^2\delta t^2 \tag{3.32}$$

the first equation is still the precursor (2.3) of the D'Alembert Equation (2.4) and is clearly an identity $0=0$ owing to $\delta x=v\delta t$; the second equation only involves explicitly m through its reciprocal Compton length. To show why, and how to implement further these equations, note that the first couple of Equations (3.32) merged together yields

$$\hbar^2\frac{\delta^2\psi}{\delta t^2}=\hbar^2v^2\frac{\delta^2\psi}{\delta x^2}, \quad \hbar^2\frac{\delta^2\psi}{\delta t^2}=\frac{(mcv)^2}{1-(v/c)^2},$$

so that one finds

$$\hbar^2\frac{\delta^2\psi}{\delta(x/c)^2}=\frac{\left(mc^2\right)^2}{1-(v/c)^2}, \quad \hbar^2\frac{\delta^2\psi}{\delta(ct)^2}=\frac{(mv)^2}{1-(v/c)^2}; \tag{3.33}$$

then the position

$$\delta t'=\frac{\delta x}{c} \tag{3.34}$$

yields by consequence

$$\hbar^2\frac{\delta^2\psi}{\delta t'^2}=\frac{\left(mc^2\right)^2}{1-\left(v/c\right)^2},\quad \hbar^2\frac{\delta^2\psi}{\delta x'^2}=\frac{\left(mv\right)^2}{1-\left(v/c\right)^2}. \tag{3.35}$$

Reasonably therefore the positions (3.30) imply (3.32), which yield (3.35) in agreement with (2.35). Hence

$$\pm\hbar^2\left(\frac{\delta\psi}{\delta t}\right)^2=\frac{\left(mc^2\right)^2}{1-\left(v/c\right)^2},\quad \pm\hbar^2\left(\frac{\delta\psi}{\delta x'}\right)^2=\frac{\left(mv\right)^2}{1-\left(v/c\right)^2}. \tag{3.36}$$

It is useful to introduce now the local limit $\delta\to\partial$ of the Equations (3.30); once more, the resulting equations take then physical meaning via this limit condition, which introduces an appropriate function $\chi=\chi\left(x,t\right)$ defined by the local properties of ψ; also now indeed the consequent position $\psi\to\chi$ turns the Equations (3.30) into the respective differential equations that represent the actual behavior of the particle. So

$$\frac{\partial^2\chi}{\partial t^2}=\pm\left(\frac{\partial\chi}{\partial t}\right)^2,\quad \frac{\partial^2\chi}{\partial x'^2}=\pm\left(\frac{\partial\chi}{\partial x'}\right)^2,\quad \chi=\chi\left(x't\right), \tag{3.37}$$

obtained equating the left hand sides of (3.35) and (3.36), are both fulfilled by

$$\chi=\mp\log\left(\xi\left(t-t_o\right)\left(x'-x_o\right)+\eta\right)+\zeta,\quad \xi=\frac{1}{\delta t_0\delta x_0}, \tag{3.38}$$

being ξ, ζ and η three arbitrary constants. The second equation remarks through the constants x_0 and t_0 that the physical dimensions of ξ are $\left(space\times time\right)^{-1}$. This equation, which emphasizes the space time range $\left(t-t_o\right)\left(x'-x_o\right)$ already found in (1.7), will be also implemented in the shortened form

$$\chi=\mp\log\left(\xi tx'+\eta\right)+\zeta. \tag{3.39}$$

An interesting corollary of (3.38) follows from $\chi=\zeta\mp\left(\log\left(\delta t/\delta t_0\right)+\log\left(\delta x/\delta x_0\right)\right)$ valid for $\eta=0$. As $\delta x/\delta x_0=m\delta x/m\delta x_0=C_0/C$, strictly speaking $C=m/\delta x$ and $C_0=m/\delta x_0=const$ are linear mass densities in the present one dimensional model; of course in a realistic four dimensional space time C and C_0 must be intended as usual mass densities, as emphasized in (2.21) and in the next (4.13). Consider χ during a fixed time range, so that $\delta t/\delta t_0$ is regarded as a time constant; then

$$\chi=const\mp\log\left(C/C_0\right),\quad const=\zeta\mp\log\left(\delta t/\delta t_0\right). \tag{3.40}$$

Although χ is dimensionless, appropriate units clarify its physical meaning: multiplying for example both sides by the energy kT already introduced preliminarily in (2.22) but to be defined shortly later, one finds

$$\chi kT=\mp\left(\mu+\mu_0\right),\quad \mu=kT\log\left(C/C_0\right),\quad \mu_0=const\,kT. \tag{3.41}$$

Hence χ is proportional to the chemical potential μ an arbitrary space time constant μ_0 apart.

Eventually note that $\delta x'$ and $\delta t'$ defined in (3.31) and (3.34) fulfill $\delta x/\delta t' = \delta x'/\delta t$; this equation is also fulfilled putting $\delta x' = \delta x/\gamma$ and $\delta t' = \gamma\delta t$ with γ arbitrary factor, in which case it reduces to identity. In particular γ could be the Lorentz factor, in fact introduced in (2.34); so one infers that

$$\delta x\delta t = \delta x'\delta t'. \tag{3.42}$$

is a relativistic invariant in different inertial reference systems. Moreover, dividing both sides by ν, write the identity

$$\frac{\delta x'}{\gamma\nu'} = \frac{\delta x}{\nu}, \quad \nu' = \frac{\nu}{\gamma}, \quad \nu' = \frac{1}{\delta t'}, \quad \nu = \frac{\gamma}{\delta t'}. \tag{3.43}$$

With γ equal in particular to the Lorentz factor $\sqrt{1-v^2/c^2}$, as suggested by δt and $\delta t'$, it is possible to regard the frequency ν' as that related to ν in different inertial reference systems R' and R. Moreover it is also possible to regard ν as the frequency recorded by an observer moving in R at rate v with respect to the frequency ν_o emitted by the source. Noting that $\nu_o c > \nu_o v$, let ν be such that $\nu_o c = \nu_o v + \nu c$ so that $\nu = \nu_o(1-v/c)$. Thus replacing in the second (3.43) one finds

$$\nu' = \nu_o \frac{1-v/c}{\gamma}.$$

This equation is nothing else but the Doppler shift of frequencies reciprocally moving at rate v along their sight line.

As (3.39) shows that both signs of (3.30) are admissible, consider now separately either sign of the Equations (3.36).

1) The negative sign yields

$$\pm i\hbar\frac{\delta\psi}{\delta t} = \varepsilon, \quad \mp i\hbar\frac{\delta\psi}{\delta x'} = p, \quad \varepsilon = \frac{mc^2}{\sqrt{1-(v/c)^2}}, \quad p = \frac{mv}{\sqrt{1-(v/c)^2}}, \tag{3.44}$$

which of course confirm (3.28); so, for $\delta \to \partial$ and thus $\psi \to \varphi$, the local limits read

$$\pm i\hbar\frac{\partial\varphi_\varepsilon}{\partial t} = \varepsilon, \quad \mp i\hbar\frac{\partial\varphi_p}{\partial x'} = p, \quad \varphi_\varepsilon = \varphi_\varepsilon(x',t), \quad \varphi_p = \varphi_p(x',t). \tag{3.45}$$

In these equations the physical meaning of $\pm p$ is immediately evident: p is actually a component of the vector p along the x axis on which is defined δx. Instead $\pm\varepsilon$ is more interesting, as it indicates the existence of states of negative energy.

Note that holds for (3.44) and (3.45) the same remark carried out for (2.3) and (2.4): also now the left hand side of (3.44) are in fact not calculable explicitly because are indeterminate not only $\delta\psi$ but also δx and δt. However are in principle calculable their limits for $\delta \to \partial$. Now also the relativistic quantities (3.45) come from and are compliant with the non-real and non-local (3.44). In effect even the Equations (3.45) bring back to the early postulates of the old quantum mechanics, despite obtained from the relativistic (3.27): this is immediately evident via the following positions:

$$\varphi=\log(\Psi),\quad \mp i\hbar\frac{\partial\varphi}{\partial x'}=\mp\frac{i\hbar}{\Psi}\frac{\partial\Psi}{\partial x'}=p,\quad \pm\frac{i\hbar}{\Psi}\frac{\partial\Psi}{\partial t}=\varepsilon,\quad \Psi=\Psi\left(x't/x_0t_0\right). \tag{3.46}$$

In this case Ψ has the same analytical form of χ. This point deserves further attention.

The relativistic equations (3.44) are implied by the invariant *xt* of (3.39), as shown in (3.42); obviously, replacing *xt* with another function $\psi_{nr}\neq\chi$, the Equation (3.37) would not hold. By consequence, in this case *p* and ε in (3.44) would be reasonably replaced by non relativistic quantities p_{nr} and ε_{nr} numerically different but having however an analogous physical meaning by dimensional reasons: the notation emphasizes the non-relativistic character of their classical approximation. Replace thus $x't$ of χ in Equation (3.39) with any function $\psi_{nr}=\psi_{nr}(x,t)$, putting for example $\psi_{nr}=\psi_x(x)\psi_t(t)$: with this Newtonian position where time and space are independent entities defining distinct dynamical variables of classical mechanics, the Equations (3.39) turn respectively into

$$\chi_{nr}=\pm\log\left(\xi\psi_{nr}+\eta\right)+\zeta,\quad \psi_{nr}=\psi_x\psi_t;$$

for simplicity of notation, the symbols of the constants have been kept unchanged. Hence the first two Equations (3.44) turn into

$$\frac{i\hbar}{\xi\psi_{nr}+\eta}\frac{\delta\psi_{nr}}{\delta t}=\pm\varepsilon_{nr},\quad \frac{i\hbar}{\xi\psi_{nr}+\eta}\frac{\delta\psi_{nr}}{\delta x}=\mp p_{nr};$$

Put eventually $\eta=0$, for example assuming η proportional to c^{-1}; as *c* is infinite in classical physics, these equations take the well known form

$$\pm i\hbar\frac{\delta\psi_{nr}}{\delta t}=\varepsilon'_{nr}\psi_{nr},\quad \mp i\hbar\frac{\delta\psi_{nr}}{\delta x}=p'_{nr}\psi_{nr},\quad p'_{nr}=\xi p_{nr},\quad \varepsilon'_{nr}=\xi\varepsilon_{nr}.$$

Clearly these expressions, suggested by the outcomes (2.23), agree with (3.28) and specify via the limit $\delta\rightarrow\partial$ which function is actually involved by the change symbol δ. Hence

$$\pm i\hbar\frac{\partial\psi_{ef}}{\partial t}=\varepsilon'\psi_{ef},\quad \mp i\hbar\frac{\partial\psi_{ef}}{\partial x}=p'\psi_{ef},\quad p'=\xi p,\quad \varepsilon'=\xi\varepsilon; \tag{3.47}$$

these results are the well known equations of the old quantum theory; the subscript "*ef*" stands for "eigenfunction". The modern quantum physics was born postulating these crucial equations, whence the importance of having found them as corollaries: the present theoretical approach brings back just to early formulation of quantum mechanics and its basic assumptions.

2) Consider now also the plus sign of (3.36), which yields

$$\hbar\frac{\delta\psi}{\delta t}=\pm\frac{mc^2}{\sqrt{1-(v/c)^2}},\quad \hbar\frac{\delta\psi}{\delta x'}=\mp\frac{mv}{\sqrt{1-(v/c)^2}}. \tag{3.48}$$

The Equations (3.64) correspond to the Equations (3.28), whereas the Equations (3.48) read $\hbar/(\pm\varepsilon\delta t)=\hbar/(\pm p\delta x)$, *i.e.* $\pm\varepsilon\delta t=\pm p\delta x$; this expression is a particular case of the Equations (3.1) regarding $\delta p = \pm p - 0$ and $\delta\varepsilon = \pm\varepsilon - 0$, where the reference boundaries of these ranges coincide with zero momentum and zero energy. Considering indeed the particular case $\pm\varepsilon_0\delta t = \pm p\delta x_0$ and subtracting side by side one finds again the expected more general result $\pm(\varepsilon - \varepsilon_0)\delta t = \pm(p - p_0)\delta x$ in agreement with (3.1) as $\delta\varepsilon = \varepsilon - \varepsilon_0$ and $\delta p = p - p_0$. So (3.48) link the operator formalism (3.28) and (3.67) to the uncertainty equations (3.1) and their relativistic implications (3.44).

Note eventually that the Equations (3.37) are well known in the operator formalism $\hat{p} = -i\hbar\partial/\partial x$, where in effect it is taken for granted that $\hat{p}^2 = -\hbar^2\partial^2/\partial x^2$; indeed (3.37) express nothing else but $\hat{p}^2 = (-i\hbar\partial/\partial x)^2 = -\hbar^2\partial^2/\partial x^2$ previously inferred from (3.30).

In conclusion this simple approach has found the operator formalism and contextually the uncertainty equation, both compatible with relativistic concepts. These outcomes have several further corollaries, the most relevant of which are shortly summarized in the following. Final remark to close this section. The range products $\delta x\delta p$ and $\delta\varepsilon\delta t$ characterize the quantum uncertainty (3.1), whereas the product $\delta x\delta t$ characterizes the invariant space time (3.38): the connection between quantum physics and relativistic physics is comprehensible corollary if space and time are mixed in either way. In this respect, what about the other mixed term $\delta\varepsilon\delta p$ also possible in alternative to $\delta x\delta p$ of (3.1)? According to (3.1) it yields

$$\delta\varepsilon\delta p = \hbar\frac{\delta\varepsilon}{\delta x} = \hbar F \tag{3.49}$$

the "new" quantity *F*, so far not explicitly concerned but only anticipated in Section 3.2 for exposition purpose only, takes in this way justification and physical meaning, it is usually known as force. The concept of pressure and energy density also follow from this result dividing both sides by the arbitrary surface Δx^2

$$\hbar\frac{\delta\varepsilon}{\delta x\Delta x^2} = \hbar\frac{F}{\Delta x^2}. \tag{3.50}$$

4. Some Classical Corollaries

Are concerned in this section several interesting outcomes still hidden in the approach hitherto outlined.

4.1. The Fermat and Maupertuis Principles

The key equations are (3.1) and (2.25). Consider an arbitrary time range Δt during which one particle moves between two coordinates x_1 and x_2 defining the total path Δx. It is possible to write

$$\Delta t = t_2 - t_1, \quad \delta\Delta t = 0;$$

the second position expresses that the time interval is arbitrary but fixed by definite time boundaries within which hold the following considerations. Since Δx traveled by the particle can be imagined as the sum of elementary ranges δx corresponding to elementary time steps δt_k, write $\Delta t = \sum\delta t_k$. Being both time and space steps arbitrary, it is possible to replace the sum with an integral and write the following chain of equations with the help of (2.25)

$$\Delta t=\int_{t_1}^{t_2}\delta t=\int_{t_1}^{t_2}\frac{n\hbar}{\delta\varepsilon}=\int_{t_1}^{t_2}\frac{\delta p\delta x}{\delta\varepsilon}=\int_{t_1}^{t_2}\frac{\delta x}{v}.$$

Hence, integrating along an element $\mathrm{d}x$ of trajectory for $\delta\to\partial$,

$$\delta\Delta t=0=\delta\int_{t_1}^{t_2}\frac{\mathrm{d}x}{v}$$

so the Fermat principle, also expressible identically as $\delta\int\mathrm{n}\mathrm{d}x=0$ with $\mathrm{n}=c/v$, is actually a straightforward corollary of the uncertainty equation.

In an analogous way one finds the Maupertuis principle. Calculate $\delta p\delta x$ for $\delta x\to\mathrm{d}x$ and $\delta\varepsilon\delta t$ for $\delta t\to\mathrm{d}t$; in this way, even considering vanishingly small range sizes still holds the concept of local velocity v_l, *i.e.* $\delta x/\delta t\to\mathrm{d}x/\mathrm{d}t=v_l$. Considering the coordinates $x_1=x(t_1)$ and $x_2=x(t_2)$ and integrating both sides, one finds according to (3.1)

$$\int_{x_1}^{x_2}\delta p\mathrm{d}x=\int_{t_1}^{t_2}\delta\varepsilon\mathrm{d}t.$$

The right hand sides involves $\delta\varepsilon_1=\delta\varepsilon(t_1)$ and $\delta\varepsilon_2=\delta\varepsilon(t_2)$. Suppose now that $\delta\varepsilon=0$ because ε is constant itself; then, being $\delta p=p_2-p_1$ by definition, one finds

$$\int_{x_1}^{x_2}\delta p\mathrm{d}x=\int_{x_1}^{x_2}p_1\mathrm{d}x-\int_{x_1}^{x_2}p_2\mathrm{d}x=0,$$

i.e. $\int_{x_1}^{x_2}p_1\mathrm{d}x=\int_{x_1}^{x_2}p_2\mathrm{d}x$ and thus $\int_{x_1}^{x_2}p\mathrm{d}x=const$ for any $p_1\le p\le p_2$. Hence, along an element $\mathrm{d}x$ of trajectory,

$$\delta\int_{x_1}^{x_2}p\mathrm{d}x=0,\quad \varepsilon=const.$$

4.2. Further Considerations on the Group Velocity

The reasoning already carried out for a beam of particles, see (2.46), is extended here considering a light beam propagating in a dispersive medium at rate $v<c$. The Equations (2.37) and (2.36) yield

$$v=\frac{\delta\varepsilon}{\delta p}=\frac{\delta\nu}{\delta\lambda^{-1}}=\frac{\delta\nu}{\delta(\mathrm{n}\nu/c)}=\frac{c}{\dfrac{\delta(\mathrm{n}\nu)}{\delta\nu}},\quad \mathrm{n}=\frac{c}{v},\quad \lambda^{-1}=\frac{\nu}{v};\tag{4.1}$$

of course v_g of a light wave packet is found through the local limit $\delta\to\partial$, *i.e.*

$$v_g=\frac{c}{\dfrac{\partial(\mathrm{n}\nu)}{\partial\nu}}.\tag{4.2}$$

It is instructive to examine closer the Equation (4.2) in order to evidence that a further aspect of the motion of a corpuscle of mass *m* is describable by a wave packet moving as a whole with at rate v_g; the reasoning involves explicitly its energy ε to describe the propagation of the overall shape of the wave packet amplitude through the space. Differentiating the Equation (2.34)

$$\delta\left(\varepsilon/c^2\right) = -\frac{p}{v^2}\delta v + \frac{\delta p}{v};$$

and replacing $p = h/\lambda$ in this equation, trivial calculations yield

$$\delta\left(\varepsilon/c^2\right) = -\frac{h}{\lambda^2 v}\left(\lambda\frac{\delta v}{v} + \delta\lambda\right).$$

Require now purposely $\delta\varepsilon = 0$, *i.e.* the wave transports a fixed amount of energy; for example $\varepsilon = const$ could be just that of one free particle. So

$$\frac{\delta\nu}{\delta\lambda} = -\frac{\nu_m}{\lambda} = -v_m, \tag{4.3}$$

being v_m the particular value of v fulfilling the given condition; the frequency ν_m is then formally implied by dimensional reasons too. Hence

$$-\delta\nu_m = -v_m\delta\lambda^{-1},$$

so that

$$v_m = \frac{\delta\nu_m}{\delta\lambda^{-1}} = \frac{\delta\nu_m}{\delta\left(\nu_m/v_m\right)} = \frac{c}{\dfrac{\delta\left(\mathrm{n}_m\nu_m\right)}{\delta\nu_m}}, \quad \mathrm{n}_m = \frac{c}{v_m};$$

then, for $\delta \to \partial$ once more, $v_m = v_g$.

The key step of the reasoning is the well defined amount energy ε transported at the rate v_g, by consequence of which results defined the frequency ν_m corresponding to the unique v_m. The different definitions of λ in (2.36) and in (4.3) are significant; their comparison yields

$$\frac{\nu_m}{\nu} = \frac{v_m v}{c^2} < 1.$$

Think now one Planck frequency (2.37) as that included in a packet of waves of different wavelengths propagating in a dispersive medium with different λ-dependent velocities: in effect, the Equations (3.1) regard $h\nu$ and h/λ as random values within energy and momentum ranges that in turn define various frequencies and momenta corresponding to $\delta\lambda$ and $\delta\nu$. Both statements agree with the fact that the propagation of the particle or its related wave correspond to v_g and not to the single phase velocities $\lambda\nu$. Just for this reason from (3.1) can be inferred the corpuscular and wave aspects of quantum physics.

The equations now obtained directly from the Equation (2.34) emphasize a new implication: neither ε_ω nor p_ω show explicit reference to the mass, which now becomes mere dimensional parameter inherent the definition of $\hbar$. Appears thus the necessity of explaining how and why the mass is apparently waived from the quantum Equation (2.37) of momentum and energy. In other words, a valid reason is required to replace m with $m - m'$, being m' a new mass even compliant with $m - m' = 0$ as a limit case. Tentatively this implies defining m as a velocity dependent variable, as in effect it has been already found in the Equation (2.33) and more specifically in (3.13). On the one hand this strategy seems at least in principle adequate to highlight why a moving mass m could turn into an immaterial wave. On the other hand further confirms should be provided next to validate the following way of describing this subtle point.

4.3. The Refractive Index

According to (3.64) and (2.35), if $m \neq 0$, then p and ε are calculable for $v < c$ only; however even $v \to c$ is admissible if contextually $m \to 0$. Implementing concurrently both limits, p and ε tend to the indeterminate forms 0/0, which admit in principle finite values. Let p' and ε' be these limit values, assumed existing by definition: the reverse question rises now, *i.e.* whether or not $v < c$ requires $m \neq 0$. The answer is negative: as the speed of photons in dispersive materials is lower than that in the vacuum, it is possible in principle that photons travel in a dispersive medium at the same v allowed to a beam of massive particles. The fact that $v < c$ is compatible with both $m = 0$ and $m \neq 0$, suggests that the kinetic mass m should actually be function of v itself: if so, then the separate correspondences $m \neq 0 \to v < c$ and $v < c \leftrightarrow (m = 0, m \neq 0)$ merge into the unique correspondence $m \neq 0 \leftrightarrow v \le c \leftrightarrow m = 0$ provided that an appropriate function of $m = m(v)$ does exist. In other words it should be true that both $m \neq 0$ and $m = 0$ are compatible with a unique $v \le c$ via $m = m(\mathrm{n})$. In fact this conclusion has been already inferred in (2.33), where the concept of mass was introduced in the present model as rest mass. The following reasoning represents the extension of this concept to the kinetic mass.

Regard m of (2.33) as a particular case of a general dynamical variable related to p through v and examine how the new concept of mass could tend to zero correspondingly to $v \to c$; is interesting in this respect the position

$$m = m'\sqrt{1 - v^2/c^2}, \quad m' \ge m, \tag{4.4}$$

which regards m as a constant mass while introducing a new mass $m' = m'(v)$. The Equation (2.33) has anticipated this conclusion in the particular case where $m = m'$ for $v \to 0$, whereas a further hint to the concept of rest mass has been provided by (3.13). Replacing formally m of (4.4) in both (3.64) one obtains $p' = m'v$ while contextually $\varepsilon' = m'c^2$; then, eliminating m' from these results, one still finds $p' = \varepsilon' v/c^2$ in agreement with (2.34). So m' fulfills the same relativistic formula of p with initial mass m, despite now the limit for $v \to c$ corresponds to the finite value $p' = \varepsilon'/c$ implemented in (3.14) and (3.15); this relationship between energy and momentum is expected in general for a wave, see Equation (2.24). Hold for m' all steps from (3.14) to (3.20). The wavelike implication of (4.4) is further acknowledged considering $\delta\varepsilon' = c\delta p'$ of (2.14).

In conclusion, according to the quantum uncertainty the behavior of a corpuscle of mass m should inherently have a wave-like propagation too, whereas the fact that $m = m'$ for $v = 0$ shows that m and m' are rest and kinetic masses. So the Equation (4.4) in fact generalizes the concepts of m and v introduced in (2.33): m' is the particular value pertinent to m at the specific speed v. As a consequence note that $p/\varepsilon = v/c^2 = p'/\varepsilon'$ define a pure number

$$\frac{\varepsilon}{pc} = \mathrm{n} = \frac{\varepsilon'}{p'c}, \quad \mathrm{n} = \frac{c}{v} \tag{4.5}$$

that introduces the refractive index of the medium where propagates an electromagnetic wave at velocity $v \le c$; owing to the Equation (6.2) in fact $\mathrm{n} \ge 1$, as it has been already introduced in (2.46) and (4.1). Moreover the position (4.4) also agrees with (2.45); indeed

$$p' = \frac{\varepsilon' v}{c^2} = \frac{h\nu' v}{c^2} = \frac{h}{\lambda}, \quad \lambda = \mathrm{n}\lambda' = \frac{c^2}{v\nu'}, \quad \lambda' = \frac{c}{\nu'}$$

takes into account that $\lambda = \mathrm{n}\lambda'$ depends on the refractive index of the medium through which propagates the electromagnetic wave or the De Broglie pilot wave. The position (4.4) introduces thus the first step to explain how and why the concept of mass does not explicitly appear in (2.45): once having introduced the refraction index, v is in fact eliminated from the equations being replaced by n. Formally this means expressing the displacement rate v of the particle in c units; yet v appears subsequently also as the rate $\lambda\nu$ at which a single wave phase propagates. This fact encourages thinking that somehow it should be possible to infer a formula that specifically emphasizes what the Equations (3.64) and (2.45) already show themselves, *i.e.* the way m and λ replace each other in defining p and ε. The mathematical approach to this task proceeds noting that

$$\delta m = m' - m = \left(1 - \sqrt{1 - v^2/c^2}\right)m' \quad \text{i.e.} \quad c^2\delta m \approx \frac{1}{2}m'v^2 + \cdots, \tag{4.6}$$

so that the right hand side represents kinetic energy. On the one hand m' corresponds to the classical mass defining the kinetic energy, although for $v \ll c$ the deviation of m' from m is irrelevant for practical purposes. On the other hand it is possible to write

$$\frac{m' - m}{m'} + \sqrt{1 - \frac{1}{\mathrm{n}^2}} = 1, \quad \left(\frac{m}{m'}\right)^2 + \left(\frac{v_g}{c}\right)^2 = 1. \tag{4.7}$$

The second equation is direct consequence of the first one; it emphasizes that the concerned velocity v is actually v_g of (2.45), because in general this latter and not v of (2.3) is related to and describable by n. This confirms that m' is the effective value of m when the particle velocity takes up just the specific value v_g pertinent to the group velocity at which propagates the wave packet. As expected $m = m'$ for $v = 0$, whereas $m = 0$ for $\mathrm{n} = 1$; in effect according to (3.13) m/m' is definable even for $v \to c$, so (2.33) and (4.4) are compliant with these limits. Regard thus the addends of (4.7) as probabilities, whose sum represents the certainty of concerning the existing particle through its mass displacement velocity or wave propagation rate. The first addend describes the probability for the particle to loose its classical kinetic mass, till to become an immaterial propagation wave; the second addend, previously introduced to express the actual velocity v_g of the particle, takes the meaning of reciprocal refraction index n of the resulting wave, being it in effect still related to the propagation rate of the wave/particle. The addends account therefore for the dual behavior of matter in a probabilistic way correspondingly to the probability of energy fluctuation, thanks to which the particle effectively displaces with velocity dependent mass or with frequency dependent propagation rate of a wave packet: indeed, one must also expect an appropriate energy fluctuation to balance the chances of mass energy loss. Obviously to this mass change correspond different p and ε and thus different λ, whence the necessity of linking m with a group of waves that spread with collective v_g given by (2.45). So the worth of (4.7) is that of having emphasized the quantum probabilistic meaning of the relativistic position (4.6).

These considerations rise however three questions.

The first one can be formulated as follows: as (4.7) is made by mass and massless terms, what determines either property of matter? Obviously the immediate answer points to the kind of experiment made on the particles constituting the body of matter. Also this is the non-real essence of quantum mechanics, which actually regards the matter neither as a packet of waves nor as a cluster of corpuscles, but as an undefined state of probabilistic mixing of both states until some experiment “creates” either state. The electron diffraction in the two slit experiment and the Thomson experiment inspired by the Millikan result elucidate the physical meaning of the addends of (4.7). To this equation is also related the physical meaning of the EPR thought paradox, showing that the quantum properties are not pre-definable outcomes according to some principles of classical mechanics, rather they are created by the experiment itself. In effect (3.1) exclude not only the concept of trajectory, but also that of distance and velocity; as shown in 3.2 the local space time coordinates must be replaced by the respective ranges, so concepts like “superluminal” distance are actually unphysical. In this sense the EPR paradox shouldn’t even be formulated: replacing systematically $x \to \delta x$ and $t \to \delta t$ are missing the concepts themselves of point to point space distance and time to time lapse needed to define any “superluminal” effect; remember that in effect according to (2.8) and (2.9) $v_0 = c$ is introduced “as such”, *i.e.* as a fundamental constant of Nature regardless the ratio δx over δt.

The second one concerns the addition of velocities. Consider an electromagnetic wave that appears in the point where $m = 0$. An example is the annihilation of m by collision with its antiparticle purposely assumed in the vacuum: one would naively expect that the new born electromagnetic wave should propagate at rate classically resulting from its own velocity c summed up to that v initially characterizing the moving center of mass of the annihilating particles. Yet (2.48) has already negatively answered this question.

The third question concerns the energy fluctuation necessary to account for the mass change when $m \to 0$. This point is concerned in the next two subsections.

4.4. Energy Fluctuation

The corpuscle/wave dualism has been accepted as compelling experimental evidence since the early experiments of electron diffraction, simply acknowledging that either behavior depends on the kind of experiment. Yet this shortcut leaves in fact unexplained why mass appears explicitly in (3.64) whereas it is hidden in the proportionality constants (2.36), despite both concern momentum and energy of a free particle. The fact that both equations have been inferred in the frame of a unique model based on the definitions (1.11) and (1.12) stimulates one to think that even this duality could find rational explanation, *i.e.* explainable by a logical physical reasoning in the conceptual frame of the present model, without need of supplementary "ad hoc" hypotheses. This hope is supported by the probabilistic character of (4.7), direct consequence of the concept of velocity dependent mass elucidated in the form (4.4): in effect the chances $m=0$ and $\lambda v=c$ or respectively $m=m'$ and $v_g=0$ appear in principle reciprocally consistent and compliant with the unity, *i.e.* the certainty that anyway something travels through the space time as amount of mass or wave: in the former case it is appropriate to think about corpuscle displacement velocity, in the latter about wave propagation rate. The validity of this idea is proposed in this subsection not only by evidencing its self-consistency, but also quoting as a verification further well known results contextually obtainable.

The results of the point 4.3 have been obtained considering initially a particle of mass m that displaces at rate v; next has been considered also its probability of mass, *i.e.* energy, fluctuation, which eventually turns it into massless electromagnetic wave or matter wave traveling at rates n^{-1} or v_g respectively. On the one hand, besides the formal similarity with the propagation of either kind of wave, the Equations (4.5) and (4.4) show that this virtual process scales both p and ε to p' and ε' by a common factor related to the refractive index. On the other hand, this also implies an energy change that occurs in a time range $\Delta t=t-t_0$, being t_0 the arbitrary time at which the mass m starts modifying its value.

Owing to (2.34), consider thus the energy change $\Delta\varepsilon=\varepsilon'-\varepsilon=p'c^2/v'-pc^2/v$ since when the particle starts loosing its initial mass m to when eventually $m\to 0$ according to (4.7). The fluctuation driven energy change is summarized by the following equations

$$\Delta\varepsilon=\frac{c^2}{v''^2}\delta\varepsilon=\mathrm{n}^2\delta\varepsilon,\quad v''^2=v'v,\quad \delta\varepsilon=p'v-v'p,\quad \mathrm{n}=\frac{c}{v''}. \tag{4.8}$$

The energy range $\Delta\varepsilon$ must not be confused with $\delta\varepsilon$ of (3.1): $\delta\varepsilon$ concerns the quantum uncertainty unavoidably constraining the arbitrary variability range allowed to the conjugate dynamical variables of any system, $\Delta\varepsilon$ is instead the specific energy fluctuation allowed in particular to the particle during the mass loss virtual process that "converts" it into a wave. The time length related to $\Delta\varepsilon$ is thus

$$\Delta t = \frac{\hbar}{\Delta\varepsilon} = \frac{\hbar}{\mathrm{n}^2\delta\varepsilon} \quad i.e. \quad \mathrm{n}^2\Delta t = \delta t. \tag{4.9}$$

The Equation (4.9) yields

$$\frac{v^2}{c^2} = \frac{\Delta t}{\delta t} = \frac{\nu_\delta}{\nu_\Delta}, \quad \nu_\delta = \frac{1}{\delta t}, \quad \nu_\Delta = \frac{1}{\Delta t} \tag{4.10}$$

these positions are easily understood; the respective energies proportional to v^2 and c^2 are also proportional to the frequencies ν_δ and ν_Δ. Of course the only way to regard this result in the wave formalism of quantum mechanics is the link between frequencies and energies, which in fact is just the Planck position: precisely in this sense the probabilistic Equation (4.7) introduces the ratio $(m/m')^2$, mass addend, and the corresponding ratio ν_δ/ν_Δ, wave addend. This confirms that the corpuscle/wave behavior has probabilistic origin and follows an energy fluctuation of quantum matter.

Are the Equation (4.8) along with its premises and implications true indeed? To support the validity of (4.7) and thus (4.8) itself, is now tested their direct consequence, the Equation (4.9), in three particular cases of major physical interest. Write first with the help of (3.1) and (2.25)

$$\Delta t = \frac{\hbar}{\mathrm{n}^2\delta\varepsilon} = \frac{v\delta x}{c^2} = \frac{\delta x^2}{c^2\delta t}, \tag{4.11}$$

noting that $c^2\Delta t$ has physical dimensions of diffusion coefficient D introduced in (2.19); this suggests that $v\delta x = \sigma D$, being σ an appropriate proportionality coefficient to be determined. So

$$\sigma D = v\delta x = \frac{\delta x^2}{\delta t} \tag{4.12}$$

The coefficient σ is crucial to specify the kind of problem precisely concerned.

1) Putting first $\sigma = 1$ means describing one particle that displaces with diffusion coefficient D through δx at average velocity *v*. Strictly speaking, as previously remarked about the Equation (3.60), in the present one space dimensional model (3.39) defines C as linear density $mass/length$ instead of the actual $mass/length^3$; yet C regarded in the usual 3-dimensional space allows to define the actual physical dimensions of flux *J* of matter, *i.e.* $mass/length^2 time$. Multiplying both sides of the first equality $v\delta x = D$ by δC, being C mass per unit volume, yields $v\delta C = D\delta C/\delta x$. This result is more appropriately rewritten as $|v|\delta C = \pm D\delta C/\delta x$; the double sign accounts for the fact that v is actually a velocity component on the *x*-axis along which is defined δx, correspondingly to the definition $v_0^2\tau$ of the Equation (2.19). Simple dimensional considerations allow defining the equation

$$J = \pm D\frac{\delta C}{\delta x}, \quad J = |v|C, \quad C = \frac{m}{V} \tag{4.13}$$

that introduces with the minus sign the concept of mass flux *J*, *i.e.* mass transferred per unit surface and time through the volume *V*; so δC is due to the diffusion driven matter transfer between the surfaces δx apart of an ideal cube of matter of volume *V*. The Equation (4.13) completes the Equations (2.19), as it is well known. Anyway, merging both expressions (4.13) of *J*, one finds $vC = -D\delta C/\delta x$; then, recalling (3.60) and (3.61) as already done in (2.22), the limit $\delta \to \partial$ yields

$$v = \mp D\frac{\partial \log C}{\partial x} = -\frac{D}{kT}\frac{\partial \mu}{\partial x}, \quad \mu = kT\log\left(\frac{C}{C_0}\right). \tag{4.14}$$

The double sign of v is obvious, being it a velocity component. For simplicity and brevity v and D have been regarded not dependent on *x*, to make quickly recognizable the link of these results with well known concepts of elementary diffusion theory; also, the diffusion process has been assumed at the constant temperature *T*. With the minus sign in (4.13), positive *D*, one acknowledges once more the definition of chemical potential μ in agreement with (3.61). Moreover, as the $-\partial\mu/\partial x$ is equivalent to a force $\mathcal{F}$, this yields also the famous Einstein-Smoluchowski relationship between mobility $\mathcal{M}$ and diffusion coefficient *D*, *i.e.*

$$D = \mathcal{M}kT, \quad \mathcal{M} = \frac{v}{\mathcal{F}}. \tag{4.15}$$

Eventually the plus sign in (4.13), which instead corresponds to negative *D*, describes phenomena like the spinodal decomposition of alloys of appropriate composition [9].

2) Putting next $\sigma = 2$ and writing thus (4.12) as $v\delta x = 2D$ means describing one particle that travels with diffusion coefficient D the distance $\delta x/2$ at average velocity $\pm v$. The factor 1/2 specifies therefore that the particle displaces around the mean coordinate $\overline{x}$ towards both sides of δx, in which case δx^2 at the right hand side of (4.11) reads $\delta x^2 \to \overline{\left(x-\overline{x}\right)^2}$ and takes thus the statistical meaning of average square displacement $\overline{\delta x^2}$ of the particle traveling through the whole range around $\overline{x}$. So the second equality (4.12) yields

$$\overline{\delta x^2} = 2D\delta t, \tag{4.16}$$

i.e. the famous Einstein equation of one dimensional Brownian motion.

3) The validity of the Equation (4.8) is further checked implementing the property $\mathrm{n} \geq 1$. Consider now a system of particles, the *i*-th of which has energy ε_i. The fact that $\Delta\varepsilon$ is in general n^2 times greater than $\delta\varepsilon = \varepsilon - \varepsilon_0$ suggests the possible chance of regarding the former as $\Delta\varepsilon = E - E_0$ and the latter as the sum of an appropriate number N of terms $\delta\varepsilon_i = \varepsilon_i - \varepsilon_{0i}$ such that $\Delta\varepsilon = \mathrm{n}^2\delta\varepsilon = \sum\left(\varepsilon_i - \varepsilon_{0i}\right)$; clearly N depends of the value of n^2 and size of all ranges $\varepsilon_i - \varepsilon_{0i}$. Anyway the initial Equation (4.8) is compatible with the position

$$E - E_0 = \mathrm{n}^2\varepsilon - \mathrm{n}^2\varepsilon_0 = \sum\left(\varepsilon_i - \varepsilon_{0i}\right) \tag{4.17}$$

simply requiring

$$E = \mathrm{n}^2 \varepsilon = \sum \varepsilon_i, \quad E_0 = \mathrm{n}^2 \varepsilon_0 = \sum \varepsilon_{0i};$$

as in principle n can take any value from 1 to ∞, the number of terms of the sum is arbitrary. The Equation (4.17) is well known and reported in all standard textbooks concerning the fluctuations of thermodynamic systems: it yields $\left(E - E_0\right)^2 = \sum\sum\left(\varepsilon_i - \varepsilon_{0i}\right)\left(\varepsilon_j - \varepsilon_{0j}\right)$ and thus $\overline{\left(\mathcal{E} - \bar{\mathcal{E}}\right)^2} = \sum\overline{\left(\varepsilon_i - \bar{\varepsilon}_i\right)^2}$ regarding appropriately E_0 and ε_{0i} as average quantities. So with $E_0 \to \bar{\mathcal{E}}$ and $\varepsilon_{0i} \to \bar{\varepsilon}$, follows then immediately

$$\overline{\left(\mathcal{E} - \bar{\mathcal{E}}\right)^2} = N\overline{\left(\varepsilon - \bar{\varepsilon}\right)^2}. \tag{4.18}$$

4.5. Liouville Theorem

An interesting question concerning (3.1) is the following: is $\delta\varepsilon = \varepsilon - \varepsilon_o$ simply an energy range or is it even compatible with the physical meaning of difference between two diverse forms ε and ε_o of energy? This question, which according to (2.25) involves $\delta p = p - p_o$ too, is answered rewriting identically (3.1) as

$$\delta\varepsilon = \frac{\delta x}{\delta t}\delta p = \delta x \frac{\delta p}{\delta t}. \tag{4.19}$$

The first equality reads

$$\frac{\delta\varepsilon}{\delta\dot{x}} = \delta p, \quad \delta\dot{x} = \frac{\delta x}{\delta t}, \tag{4.20}$$

the second equality reads

$$\frac{\delta\varepsilon}{\delta x} = \delta\dot{p}, \quad \delta\dot{p} = \frac{\delta p}{\delta t}. \tag{4.21}$$

Now fulfill the idea that $\delta\varepsilon$ defines the difference of two distinct energies, specifically T and U introduced in section 2.3, which implies the chance of writing in general $\varepsilon = \mathrm{T} \pm \mathrm{U}$. To highlight this point, concerning in particular the energies already introduced in (2.41) and (2.42), introduce the following positions

$$\delta\varepsilon = \delta\mathrm{T} \pm \delta\mathrm{U}, \quad \mathrm{T} = \mathrm{T}\left(\dot{x}, x\right), \quad \mathrm{U} = \mathrm{U}\left(x\right) \tag{4.22}$$

in this way the sign of δp in (4.20) is uniquely defined since T only depends on $\dot{x}$, whereas is expected the double sign in (4.21) because both energies T and U depend on *x*. As in effect δp is the component of $\delta\boldsymbol{p}$ along the *x*-axis, so that it can actually take in principle both signs, rewrite explicitly (4.21) as

$$\frac{\delta\varepsilon}{\delta x} = \pm\delta\dot{p} = \delta\dot{p}' - \delta\dot{p}'_o, \quad \delta\dot{p}' = \frac{p - p'_o}{\delta t}, \quad \delta\dot{p}'_o = \frac{p_o - p'_o}{\delta t} \tag{4.23}$$

the double sign on the one hand emphasizes that both *p* and p_o are actually components of the vectors $\boldsymbol{p}$ and $\boldsymbol{p}_o$ along the *x*-axis. The last two equations also agree with the fact that in principle

$$\delta\dot{p}' \gtrless \delta\dot{p}'_o$$

in lack of any information about the ranges, both inequalities are actually possible. Regarded in this way, *i.e.* implementing range boundaries arbitrary and independent each other, the notation (4.23) effectively defines $\delta\varepsilon$ as difference of two energies reasonably dissimilar according to (4.22). Taking the ratios side by side of the first Eqs (4.20) and (4.21) one finds

$$\frac{\delta\dot{x}}{\delta x}=\pm\frac{\delta\dot{p}}{\delta p} \tag{4.24}$$

It is immediate to link (4.24) and (4.22), noting that the former defines at both sides ratios with physical dimensions of reciprocal time range. Multiplying both sides by $\hbar$, the equations

$$\hbar\frac{\delta\dot{x}}{\delta x}+\hbar\frac{\delta\dot{p}}{\delta p}=0,\quad \hbar\frac{\delta\dot{x}}{\delta x}-\hbar\frac{\delta\dot{p}}{\delta p}=0 \tag{4.25}$$

define energies that, in agreement with (2.42) and (2.38), correspond respectively to

$$\delta\mathrm{H}=\delta\mathrm{T}+\delta\mathrm{U}=0,\quad \delta\phi=\delta\mathrm{T}-\delta\mathrm{U}=0.$$

Hence simple considerations on the range boundaries imply the concepts of Hamiltonian and Lagrangian according to the previous Equations (2.38) and (2.42): ϕ has been identified with the Lagrangian of a particle, H with the Hamiltonian of the system. In particular, is of interest here

$$\mathrm{H}=\mathrm{T}+\mathrm{U}=const,\quad \frac{\delta\dot{x}}{\delta x}=-\frac{\delta\dot{p}}{\delta p}$$

for the following reason. According to the quantum uncertainty, the left hand side of (3.1) reads $\delta\boldsymbol{x}\cdot\delta\boldsymbol{p}=n\hbar$, where the number of scalars, so far intuitively associated to the three usual space dimensions only, is actually arbitrary, *i.e.* extensible to any number j of extra-dimensions required by some theories or, alternatively, to the number of freedom degrees allowed to the system of particles: in fact any freedom degree has its pertinent $\delta\boldsymbol{x}$ and $\delta\boldsymbol{p}$. Thus it is sensible to introduce the dimensionless quantity $(\delta x\delta p/\hbar)^j$ where fall all points in the multidimensional phase space defined by the sizes of all δx and δp of the corresponding particles with respect to $\hbar$. Accordingly

$$\delta\Omega=(\delta p\delta x/\hbar)^j$$

yields

$$\frac{\delta(\delta\Omega)}{\delta t}=j\frac{(\delta p\delta x/\hbar)^j}{\delta x}\delta\dot{x}+j\frac{(\delta x\delta p/\hbar)^j}{\delta p}\delta\dot{p}=j\left(\frac{\delta\Omega}{\delta x}\delta\dot{x}+\frac{\delta\Omega}{\delta p}\delta\dot{p}\right)$$

the range $\delta\Omega$ includes all points of coordinates x and p falling within $(\delta x\delta p/\hbar)^j$ elementary cells of j-dimensional volume $(\delta x\delta p)^j$ in the phase space. So

$$\frac{1}{j}\frac{\delta(\delta\Omega)}{\delta t}=\frac{\delta\Omega}{\delta x}\delta\dot{x}+\frac{\delta\Omega}{\delta p}\delta\dot{p}$$

yields then

$$\frac{1}{j}\frac{\delta(\delta\Omega)}{\delta t}=\delta\Omega\left(\frac{\delta\dot{x}}{\delta x}+\frac{\delta\dot{p}}{\delta p}\right)=0$$

according to the first (4.25) $\delta(\delta\Omega)/\delta t=0$, *i.e.* the volume $\delta\Omega=const$ along phase space trajectories where $\mathrm{H}=const$.

5. Some Thermodynamic Corollaries

The last results have somehow linked the relativistic Equation (2.25) to important results of classical statistical thermodynamics. The importance of this topic is shortly highlighted in the following three subsections.

5.1. Statistical Sets of Particles

Let us implement once more the Equation (2.26), and calculate the change of $v/c=pc/\varepsilon$ according to the following chain of equations

$$\delta\left(\frac{v}{c}\right)=\frac{\delta(pc)}{\varepsilon}-\frac{pc}{\varepsilon^2}\delta\varepsilon=\frac{pc}{\varepsilon}\delta\log\left(\frac{pc}{\varepsilon}\right).$$

Since by definition

$$\delta\log\left(\frac{pc}{\varepsilon}\right)=\log\left(\frac{pc}{\varepsilon}\right)-\log\left(\frac{p_oc}{\varepsilon_o}\right),$$

being p_o and ε_o arbitrary constants, it is possible to write

$$\delta\left(\frac{v}{c}\right)=\frac{v}{c}\log\frac{v}{c}-\frac{v}{c}\log\frac{v_o}{c}.$$

Consider now preliminarily the case of an ideal gas of non-interacting free particles/atoms/ions/molecules and let p_i and ε_i the momenta and energies of each particle. Then, owing to the last equality, it is possible to write for each *i*-th particle

$$\delta\Pi_i=\Pi_i\log\Pi_i-\Pi_i\log\Pi_o,\quad \Pi_o=\frac{v_o}{c},\quad \Pi_i=\frac{v_i}{c}<1;$$

moreover it is also possible to sum terms like this of each particle over all particles of the system, so that it is possible to write

$$\sum_i\delta\Pi_i=\sum_i\Pi_i\log\Pi_i-\log\Pi_o\sum_i\Pi_i$$

whence

$$N\overline{\delta\Pi}=\sum_i\Pi_i\log\Pi_i-N\bar{\Pi}\log\Pi_o,\quad \overline{\delta\Pi}=\frac{1}{N}\sum_i\delta\Pi_i,\quad \bar{\Pi}=\frac{1}{N}\sum_i\Pi_i,\qquad (5.1)$$

being N the number of particles of the system. Note that this result is actually more general than prospected here. Suppose first two interacting particles only; in this case we expect p_1' and ε_1' for the first particle and p_2' and ε_2' for the second one because of their interaction: despite the first (2.34) holds for a free particle, it is reasonable to think that changing appropriately $p_i\to p_i'$ and $\varepsilon_i\to\varepsilon_i'$ one can describe at least approximately even an interacting particle.

For example it is possible to replace p_i/ε_i with $p'_i/\varepsilon'_i = \sigma_i\left(p_i/\varepsilon_i\right)$, being σ_i an appropriate correction factor. Anyway, p'_1/ε'_1 and p'_2/ε'_2 are in principle calculable; summing these terms, the left hand side of (5.1) involves $v'_1/c + v'_2/c$. In the case of three particles mutually interacting one would obtain $p''_1/\varepsilon''_1 + p''_2/\varepsilon''_2 + p''_3/\varepsilon''_3$ defining $v''_1/c + v''_2/c + v''_3/c$, and so on for any number of particles all mutually interacting. On the one hand this means that now the previous Π_i is replaced by $\Pi'_i = \sigma_i v_i c$, whereas the summation is possibly extended to a different number of terms. On the other hand this reasoning holds also for $\Pi''_i = \sigma'_i v'_i c$ and also for multiple primed probabilities. In fact, summing all v_i/c or all v''_i/c does not change the conceptual statistical meaning of the sum; in other words, whatever v''_i might be, one could include appropriate correction factors to the various v_i of the allowed states; normalizing the sums, one still obtains an equation like (5.1). To calculate how each v_i turns into v'_i and next into v''_i because of these interactions, is in general difficult and must be examined case by case; yet, if we content ourselves to describe the evolution of the system as a whole from the non-interacting to the interacting state, the form of the final equation is still similar to the previous (5.1) obtained for free particles. Omitting for simplicity the primed or multiple primed notations for Π_i and v_i in the following, introduce the positions

$$S = N\sigma S_o - N\overline{\delta\Pi}, \quad S_o = -\Pi_o \log\Pi_o, \quad S = -\sum_i \Pi_i \log\Pi_j, \quad \overline{\Pi} = \sigma\Pi_o, \quad (5.2)$$

where σ is now a proportionality constant. The factor N in (5.2) simply shows that S is an extensive property. Introduce the condition expressed by the equivalent positions

$$\overline{v} = \overline{\Pi} = const, \quad \overline{\delta v} = \overline{\delta\Pi}, \quad (5.3)$$

where $const$ can be 0 or more in general $\neq 0$ in this one dimensional model where v is actually a component of $\boldsymbol{v}$ that can take both signs. The first condition regards a completely disordered system of particles regarded as a whole at the equilibrium, whose velocities are randomly distributed both by modulus and direction with equal probability. The second condition assumes a macroscopic system in an unstable situation out of equilibrium, e.g. gas with an internal pressure gradient due to a non uniform distribution of velocities; this can happen for example for a system of charged particles in an inhomogeneous external field. Whatever its particle velocity distribution might be, both chances are assumed compatible with the third position (5.2).

Actually nothing compels these positions, which in effect are purposely introduced to plug the present considerations into the realm of statistical mechanics. In practice $\delta\left(velocity\ distribution\right) = 0$ shows that the equilibrium corresponds to the maximum possible disorder of the system as concerns the velocity distribution of its constituting particles. This statement, assumed valid in general and not in the present one dimensional case only, can be regarded as boundary condition of (5.2) as it implies $\Pi = const$, in this particular case $S = const$. According to (5.2) this constant can be nothing else but the right hand side, *i.e.*

$$S_B = -\sigma' \log \Pi_o, \quad \sigma' = \sigma\Pi_o \tag{5.4}$$

This is nothing else but the Boltzmann definition of dimensionless entropy: note that Π_o constant indeed does not mean that it has one fixed value only, but that it does not depend on the index of summation states *I*, whereas it depends of course on *T*.

This simple procedure has introduced the function *S* as sum of consistent functions $-\Pi_i \log \Pi_i$ of all particles of the system; the summation over *i* has been extended to the velocities v_i of all particles of the system. This summation is surely positive and finite because all $v_i < c$.

It is possible to ask at this point whether this kind of equation is uniquely referable to the property $\sum_i \Pi_i = const$, or it has a more general worth, e.g. in the case of probability distribution function of states such that $\sum_i \Pi_i = 1$. This expectation is sensible, being a particular case of the second (5.3). The next subsection concerns just this point.

5.2. The Entropy

The starting point is now (3.38) with the minus sign. The way to implement this equation is similar to that just described for the Equation (5.1): any space time factor $\delta x \delta t$ is regarded as $\delta x_j \delta t_k$, with notation that goes back to the section 1 in order to specify an arbitrary *j*-th state of a system of particles at the time t_k. The system defined in this way is a statistical set in the sense previously highlighted for each $v/c \to v_i/c$, in agreement with the definition of $v_j = \delta x_j / \delta t_k$: in other words, the variation of configuration of the system implies reasonably the change of local space coordinates of a cluster of *j*-th particles enclosed in δx_j during the time range δt_k: both ranges define a possible state in the phase space as described in the subsection 4.5. Actually both δx and δt were inherent the definition of v_i in subsection 5.1; similarly *W* must be be introduced here in order to describe the non-instantaneous evolution of a local small volume of the system during space time ranges that represent its configuration change rate. The Equations (3.60) and (3.61) show that this way of thinking allowed to infer the chemical potential μ hidden in χ; let us examine here the possibility of extracting further thermodynamic information from this function.

The algebraic steps are listed one by one after rewriting (3.38) as

$$\chi_- = \zeta - \log W, \quad W = \xi \delta x \delta t + \eta. \tag{5.5}$$

1) On the basis of the section 5.1, define

$$u = a_k W \chi_- = E + a_k S, \quad E = \zeta a_k W, \quad S = -W \log W, \quad a_k = a_k(t_k) > 0, \tag{5.6}$$

being a_k positive factor dependent on the time t_k only;

2) regard $\delta x \delta t \to \delta x_j \delta t_k$, *i.e.* any local space time coordinate *xt* is defined as one that characterizes the *j*-th state of each particle in the space range δx_j during the time range δt_k, which implies $W \to W_{jk}$ while $u \to u_{jk}$ and $S \to S_{jk}$ as well

$$W_{jk}=W_{jk}\left(\delta x_j\delta t_k\right),\quad S_{jk}=S_{jk}\left(\delta x_j\delta t_k\right),\quad u_{jk}=u_{jk}\left(\delta x_j\delta t_k\right)\tag{5.7}$$

3) sum over all allowed states *j* accessible during an assigned δt_k by all particles in the phase space

$$U_k=\zeta a_k\Theta_k+a_kS_k,\quad U_k=\sum_j u_{jk},\quad \Theta_k=\sum_j W_{jk},\quad S_k=-\sum_j W_{jk}\log W_{jk},\tag{5.8}$$

whereas the factor a_k is defined by

$$U_k=U_o+a_kS_k,\quad U_o=\zeta a_k\Theta_k;\tag{5.9}$$

4) the last Equation (5.9) defines a "new" quantity *T* called temperature

$$U_o=U_k-TS_k,\quad a_k=\frac{\delta U_k}{\delta S_k}=T,\quad T=T\left(t_k\right)\tag{5.10}$$

uniquely defined for a body of matter at the thermal equilibrium. Note that the first equation has been written introducing at the left hand side the summation over W_{jk} only. Also note that

$$\frac{\delta S_k}{\delta W_{jk}}=\frac{\delta S_{jk}}{\delta W_{jk}}=-\log W_{jk}-1$$

i.e.

$$-\log W-\frac{\delta S_{jk}}{\delta W_{jk}}=1\tag{5.11}$$

the *j*-th addend contributing to S_k is to be considered to calculate the right hand side. Multiplying both sides by W_{jk} one finds

$$W_{jk}\frac{\delta S_k}{\delta W_{jk}}=-W_{jk}\log W_{jk}-W_{jk}$$

and then, summing over *j*, owing to (5.8) one finds

$$\Theta_k=S_k^o+S_k,\quad S_k^o=-\sum_j\left(W_{jk}\frac{\delta S_k}{\delta W_{jk}}\right).\tag{5.12}$$

Normalizing via Θ_k, this result reads

$$1=\frac{S_k^o}{\Theta_k}+\frac{S_k}{\Theta_k}\tag{5.13}$$

If $S_k^o>0$ this equation emphasizes the certainty resulting from the sum of two positive terms, which therefore can be regarded as probabilities. If S_k measures the disorder of the system, then reasonably S_k^o measures the order: the sum of these probabilities yields the certainty that both order and disorder concur to define the state of any system. In other words, any system can be partially ordered and partially disordered; e.g. some parts of a crystal lattice can contain in general local point and/or line pile up defects inside a surrounding defect free volume.

This probabilistic interpretation is possible if $S_k^o\geq 0$: in other words, the probability of modifying the local order/disorder of the system requires according to (5.12) $\delta S_{jk}/\delta W_{jk} \le 0$ inside any δx_j at different δt_k at which is calculated δS_{jk}. Let be therefore $\delta S_{jk} = S_{jk+1} - S_{jk}$ and $\delta W_{jk} = W_{jk+1} - W_{jk}$ the changes allowed to occur within any space range δx_{jk} at any time within δt_k and rewrite W of (5.5) according to the positions (5.7) via (3.1) and (2.28); once more the space ranges, and not the local space time coordinates they represent, are physically appropriate to describe the changes in the system. Replacing $\delta t = \hbar/\delta\varepsilon$ and $\delta x = \hbar/\delta p = v\hbar/\delta\varepsilon$ one finds

$$W_{jk} = \xi \frac{\hbar^2}{\delta\varepsilon^2} \frac{\delta x_{jk}}{\delta t_k} + \eta$$

having expressed $v = \delta x_{jk}/\delta t_k$ according to the current notation; clearly $\delta\varepsilon_{jk}$ is the pertinent energy change corresponding to the configuration change in progress within δx_{jk}. Hence, keeping $\delta\varepsilon_{jk}$ and δx_{jk} constants, write

$$\delta W_{jk} = -\xi \frac{\hbar^2}{\delta\varepsilon_{jk}^2} \delta x_{jk} \left(\frac{1}{\delta t_{k+1}} - \frac{1}{\delta t_k} \right)$$

and thus

$$\frac{\delta S_{jk}}{\delta W_{jk}} = -\left(\xi \frac{\hbar^2}{\delta\varepsilon_{jk}^2} \delta x_{jk} \right)^{-1} \frac{S(\delta t_{k+1}) - S(\delta t_k)}{\delta t_{k+1} - \delta t_k}.$$

In this way S_k and S_k^o describe the changes occurring in δt_{k+1} with respect to δt_k in the given region δx_{jk} of the system. Certainly the local S_{jk} is due to the corresponding local changes of x_{jk} and ε_{jk}; however it is in principle possible that even at δt_{k+1} both these latter remain still included in the same range size δx_{jk} where they were at δt_k; this simply means that δS_{jk} is small enough to imply correspondingly small changes of x_{jk} and ε_{jk} that therefore still remain included within the same δx_{jk}. While acknowledging that this is in principle admissible because all range sizes are in principle arbitrary, it is interesting to compare what happens at δt_{k+1} and δt_k. If for example $\delta t_{k+1} > \delta t_k$ by definition, *i.e.* the former is greater than the latter because it must include increasing values of local time coordinates t_k, then $\delta W_{jk} < 0$ implies $W_{jk} < W_{jk+1}$. The negative sign of δW_{jk} means that on the one hand $S_k^o \ge 0$ fulfils via $\delta S_{jk}/\delta W_{jk} \le 0$ the probabilistic meaning of (5.13) and that on the other hand it also implies all $\delta S_{jk} \ge 0$. Thus summing over *j* all terms δS_{jk} at all δt_k one infers

$$\delta S_k = \sum_j \delta S_{jk} \ge 0. \tag{5.14}$$

Clearly this is just the second law of thermodynamics because, as written, it concerns an isolated system; the conclusion is in effect true if no external action perturbs the system. If not so, then any action altering substantially the configuration of the system modifies by consequence the *j*-th range size too; in general different δx_{jk} and $\delta\varepsilon_{jk}$ are reasonablyimplied before and after the external action. Thus, in particular, it can result that $\delta x_{jk+1} < \delta x_{jk}$ while however (5.13) can be again fulfilled: $\delta S_{jk}/\delta W_{jk} \le 0$ still holds even with $\delta W_{jk} > 0$ but $\delta S_k < 0$. Clearly in the system no longer isolate the external action has modified the spontaneous tendency towards increasing entropy.

It is worth remarking once more that the evolution of the physical system has implemented two subsequent time lapses δt_k and δt_{k+1}, not two deterministic time coordinates t_k and t_{k+1}: these latter and the respective deterministic ε_{jk} and ε_{jk+1} representing the external action would be incompatible with the Heisenberg principle.

5.3. The Statistical Distributions

Let the change δW of W be $\delta W = W \pm w$, being w an arbitrary amount added or subtracted to the initial value of W. On the one hand W can increase or decrease by any physical reason with respect to its initial value; the double sign indicates that no reason is guessable to expect that the change consists of either increase only or decrease only of the initial value W. On the other hand it is also reasonable to expect that $\delta W = qW_0$, being q an arbitrary proportionality factor and W_0 an arbitrary value allowed to W consistent with (5.5); this position means that anyway the change δW implies a new quantity still related to the meaning of thermodynamic probability W_0 coherent with W. In other words, W_0 is such that $W' = W + qW_0$ and $W'' = W - qW_0$ are respectively compatible and physically consistent with $W \pm w$. The fact that both W' and W'' must fulfill (5.8) likewise the initial W, allows expecting the consistency of the following considerations with the equations up to (5.10) as well. If so, then

$$\delta W = W \pm w_\pm = q_\pm W_0, \quad q_\pm > 0 \tag{5.15}$$

yields

$$\frac{1}{q_\pm} = \frac{W_0/w_\pm}{W/w_\pm \pm 1}; \tag{5.16}$$

put in this form, once more the space time Equation (5.5) of W is implemented via W/w and related W_0/w, similarly to the position (5.8) leading to the results (5.10). The Equation (5.15) has been written in order to emphasize how w is to be regarded in agreement with either sign, *i.e.* $W + w_+ = q_+ W_0$ and $W - w_- = q_- W_0$. In conclusion, recalling the Equation (5.5),

$$\frac{w_\pm}{q_\pm} = \frac{W_0}{\exp\left(\chi_- - \zeta - \zeta_0\right) \pm 1}, \quad w_\pm = \exp\left(\zeta_0\right);$$

also here appears the space time function χ_-. Hence, according to the reasoning to infer the Equation (4.14) via (3.60) and (3.61), at any given time $\chi_- - \zeta = \log\left(C/C_0\right) + const$ and thus $\mu/kT - \mu_0/kT$. In conclusion

$$\frac{w_\pm}{q_\pm} = \frac{W_0}{\exp\left(\left(\mu - \mu_0\right)/kT\right) \pm 1}. \tag{5.17}$$

This equation follows from the arbitrariness of μ_0, consistent with that of ζ_0; the multiplicative factor w^{-1} has simply included ζ_0 in the constant addend of chemical potential μ together with ζ. Implement now either

$$W = q_- W_0 - w_+, \quad W = q_+ W_0 + w_-$$

of (5.15): being $W > 0$ by definition, there is no constrain to the number w_-

related to the negative sign in the Equation (5.17), whereas the positive sign of this equation requires $w_+ \le q_- W_0$ *i.e.* $w_+ / q_- W_0 \le 1$. This constrain suggests the possible physical meaning of $w_\pm / q_\pm$ in (5.17). Let W_0 be the numbers of states with a given energy μ, *i.e.* the degeneracy of the state, and $w_\pm / q_\pm = N_\mu$ the number of particles in the given state; if so, then it is easy to realize that in the latter case N_μ can take only the values 0 and 1 whatever μ might be [10]. All details published elsewhere are omitted here for sake of brevity. Thus (5.17) is the well known formula of statistical distribution of fermions and bosons with degeneracy W_0.

5.4. The Phase Space

Entropy and Liouville theorem, both previously inferred, are the key concepts to introduce the phase space. As this topic is well known, are reported here just a few remarks aimed only to emphasize the link between space time and phase space; *i.e.* the concept of space time is actually the third essential ingredient to introduce "*ab initio*" the statistical mechanics. To this purpose consider in particular the Equations (3.39) and (3.1).

Being *x* and *t* arbitrary and independent variables, which represent for example the space coordinate of a given particle at various times in the space time, any value of *xt* can be obtained keeping constant either factor and allowing appropriate values of the other one; both ways of defining an arbitrary space time coordinate $x_j t_k$ are numerically and conceptually equivalent to describe each one among *N* particles of the system at given time t_k in the range of space coordinates $x_{j1} \le x_j \le x_{jN}$ or at x_j during the time range $t_1 \le t_k \le t_2$. According to (3.1), indeed, the space time coordinate of each particle is defined within allowed variability ranges $\delta x_j = x_{jN} - x_{j1}$ and $\delta t_k = t_{k2} - t_{k1}$. So is physically significant the amount $\delta x_j \delta t_k = \hbar^2 / \delta p_j \delta \varepsilon_k$, whatever these range sizes might be. To highlight this point consider the following equations obtained implementing (3.1) and (2.28)

$$\delta x_j \delta t_k = \frac{\hbar^2}{\delta p_j \delta \varepsilon_k} = \frac{\hbar^2}{\left(\delta \varepsilon_j / v_j\right)\left(v_k \delta p_k\right)},$$

whence

$$\frac{\delta x_j}{v_j} v_k \delta t_k = \frac{\hbar^2}{\delta \varepsilon_j \delta p_k} = \delta t_j \delta x_k$$

and thus, comparing the initial and final ranges of coordinates,

$$\delta x_j \delta t_k = \delta x_k \delta t_j \tag{5.18}$$

the initial equation regards the *j*-th space coordinate x_j at the time t_k, the former defined within the interval δx_j the latter in the time range δt_k; the final equation rewrites the first one with exchanged indexes *j* and *k*. As in effect the first two equations are summarized in the third one, it means that the concerned particle is described at different times t_j and t_k by different space ranges δx_j and δx_k, to which correspond the respective momenta δp_j and

δp_k; moreover it also follows $\delta p_j \delta\varepsilon_k = \delta p_k \delta\varepsilon_j$, as it must be because anyway the Equation (3.1) must be fulfilled no matter how any particle moves in the space time. In other words, as *j* and *k* are not specified or specifiable, the particle moves actually through any random space and time ranges in the phase space according to its position and momentum of the space time. This can be better evidenced and generalized rewriting with trivial manipulations the last equation

$$\begin{gathered}\delta x_j \delta t_k = \delta x_j \frac{\hbar}{\delta\varepsilon_k} = \delta x_j \frac{\delta p_q \delta x_q}{\delta\varepsilon_k} = \frac{\delta x_q \delta p_q}{F_{kj}} = \delta x_q \delta t_{kjq} \\ F_{kj} = \frac{\delta\varepsilon_k}{\delta x_j}, \quad \delta t_{kjq} = \frac{\delta p_q}{F_{kj}},\end{gathered} \tag{5.19}$$

i.e. one particle initially at any random x_j within δx_j at the random time t_k included in δt_k is actually found within another δx_q at the subsequent time t_{kjq} within time range δt_{kjq}. Obviously this chain could be further extended starting again from the last term $\delta x_q \delta t_{kjq}$, which in effect has the same form of the first one but is simply rewritten with different subscripts; the first and last terms of this chain represent different space time coordinates and thus its ability of the particle to fill various coordinate and momentum ranges defining the whole phase space. All accessible local coordinates of space time correspond to the respective local coordinates of space phase, with equal probability.

5.5. Further Comments about the Diffusion Coefficient

The diffusion coefficient *D* introduced in (2.19) is usually concerned in problems of matter displacement under non-equilibrium conditions, essentially due to concentration gradients; the same holds for the heat diffusion coefficient (2.22) in non-thermal equilibrium problems, typically in the presence of temperature gradients. However, the four equations from (4.13) to (4.16), as well as the next (5.20) and (5.21), suggest a more profound physical meaning of *D*. In this respect deserve attention the following three remarks.

1) The dimensional definition of *D* is $\hbar/m$; this yields $\delta D/D = -\delta m/m$, *i.e.* $\delta\log(D/D_o) = -\delta\log(m/m_o)$, being of course D_o and m_o arbitrary constants. Then, reasoning likewise in (4.14), the right hand side yields $\delta\log((m/V)/(m_o/V)) = \delta\log(C/C_o)$, being C the amount of mass in a given volume *V*. Hence, being $\log(C/C_o) = \mu/kT$, as found in (3.60) and (3.61), one finds $\delta\log(D/D_o) = -\delta(\mu/kT)$ and therefore

$$\frac{D}{D_o} = \exp(-\mu/kT) \tag{5.20}$$

this is the usual form to express the dependence of diffusion coefficient on temperature via the activation energy μ and the reference constant D_o.

2) Assume now a body of matter of mass *m* in equilibrium at temperature T and implement the reasonable idea that both *D* and $\mathcal{M}$ take finite ranges of allowed values. Let $D_{\min}$ and $\mathcal{M}_{\max}$ be the respective limit values of interest here; is then significant the particular case where the Equation (4.15) concerns

the minimum temperature $T_{\min}$ defined as follows

$$T_{\min}=\frac{D_{\min}}{k\mathcal{M}_{\max}}=\frac{\epsilon_{zp}}{k}. \tag{5.21}$$

Dimensional considerations are useful to guess an order of magnitude estimate of $T_{\min}$. The reciprocal mobility $\mathcal{M}^{-1}$ has physical dimensions $mass/time$, whereas D is $h/mass$; so their product represents the minimum energy ϵ_{zp}. The notation emphasizes that the energy of interest to calculate $T_{\min}$ excludes the contribution of thermal vibrations, being instead due to the mere confinement of a particle or a body of matter within a finite delocalization range δx; accordingly, it is sensible to define $T_{\min}$ as ε_{zp}/k. In fact the Equations (3.1) justify the existence of this form of energy and related force.

Consider indeed one particle of mass m ideally delocalized between two infinite potential walls δx apart; in a one dimensional model it is possible to write

$$\varepsilon_{zp}=\frac{\delta p^2}{2m}=\hbar\frac{\omega}{2},\quad \omega=\frac{\hbar}{m\delta x^2}=\frac{\sigma D}{\delta x^2}, \tag{5.22}$$

having expressed $D=\sigma\hbar/m$ via an appropriate proportionality constant σ. This result is understandable thinking an oscillating particle confined in δx, so that m bounces back and forth between the potential walls with frequency $\omega=1/\delta t$. In fact δt is the time lapse to complete one oscillation cycle; $\delta p=p_{\max}-0$ is the range defined by the maximum delocalization momentum $p=p_{\max}$ related to the range size and $p=0$ when the particle inverts its motion on both potential walls. This picture agrees with $\delta p\neq 0$, *i.e.* with the physical impossibility of conceiving a localized particle at rest and thus with $p=0$ in fixed point exactly defined. The circular frequency here introduced is justifiable from a more realistic three dimensional point of view, where the back and forth one dimensional motion of m reads actually $\hbar=m\omega\delta x^2$ and thus $\epsilon_{zp}=m(\omega\delta x)^2/2$; *i.e.* m describes a closed circular path at tangential velocity $v_{\tan}=\omega\delta x$ inside its confinement delocalization volume V, so far not yet introduced explicitly. In effect it is also possible to evidence the confinement volume writing

$$\epsilon_{zp}\approx\frac{\delta p_{zp}^2}{2m}=\frac{\hbar^2}{2m\delta x_{zp}^2}=\frac{\hbar^2}{2mV_{zp}^{2/3}},\quad F_{zp}\approx\frac{\hbar^2}{2mV_{zp}},\quad V_{zp}=\delta x_{zp}^3. \tag{5.23}$$

Anyway it is sensible that $T_{\min}$, being presumably a fixed value, cannot depend on the arbitrary m and specific V_{zp}; rather $T_{\min}$ is to be regarded as a universal property of matter uniquely defined. Both requirements suggest restarting from the relativistic energy equation $\varepsilon=pc^2/v$ of one free particle of arbitrary mass m, which however must no longer appear explicitly. Implementing thus the wave expression of momentum $p=h/\lambda$, which in fact allows introducing the expected oscillation behavior as that related to the concept of wavelength λ, one finds

$$\varepsilon=\frac{pc^2}{v}=\frac{hc^2}{\lambda v}, \tag{5.24}$$

which thus also defines

$$\epsilon_{zp}^{\min}=\frac{hc^2}{\left(\lambda_{zp}v\right)_{\max}}.$$

The 3D generalization of this result is obtained imagining an arbitrary amount of mass delocalized in an appropriate range $2\Delta r$, regarded as the diametric size of a hypersphere of radius Δr to which is related the maximum value of λ_{zp}: the idea is to implement steady matter waves of wavelengths λ_{zp} propagating through the hypersphere at rate $v_{\max}$. Regard thus an arbitrary mass of an isolate free corpuscle ideally bouncing within one diametric distance, whose extent corresponds to one half wavelength; the largest zero point wavelength is that with steady nodes on the opposite boundaries of the hypersphere diameter and is thus $\lambda_{zp}=4\Delta r$. This implies that $\epsilon_{zp}^{\min}$ defined by the zero point momentum wavelength λ_{zp} corresponds to back and forth delocalization through twice the diametric size $4\Delta r$ of the hypersphere: this is the physical meaning of (5.22) where $\nu_{zp}=v/2\left(2\Delta r\right)$ whence $h\nu_{zp}=\hbar\omega_{zp}$. Also, imagining asymptotically $v\to c$ to simulate $v_{\max}$ and recalling (5.22), it is possible to conclude

$$\epsilon_{zp}^{\min}=\frac{hc}{4\Delta r},\quad T_{\min}=\frac{\epsilon_{zp}^{\min}}{k}=\frac{hc}{4k\Delta r},\quad \omega=\frac{2\epsilon_{zp}^{\min}}{\hbar}=\frac{\pi c}{\Delta r}. \tag{5.25}$$

These results will be calculated later; regardless of the numerical values, however, it is possible to remark since now some interesting implications:

- The Nernst theorem is automatically fulfilled, *i.e.* the absolute zero actually does not exist being clearly impossible to remove the zero point energy, which is an intrinsic feature itself of any amount of confined matter.
- As expected, the related zero point temperature $\epsilon_{zp}^{\min}/k$ does not depend on the specific amount and physical nature of *m*.
- Is in principle possible the quantization of temperature, which accordingly should start from $T_{\min}$ and change by discrete steps of the order of $T_{\min}$ itself.

3) The reasoning to infer (5.21) and (5.25) introduces $D_{\min}$, whose physical meaning is relevant: it implies that $T_{\min}$ is somehow linked to the possible granular structure of the space time.

To show this last point, calculate the change $\delta\left(x_j t_k\right)$ of the space time coordinates around any local coordinate x_j and t_k. The result is elucidated by the following chains of equations implementing once more the Equations (3.1):

$$\delta\left(x_j t_k\right)=t_k\delta x_j+x_j\delta t_k=\left(t_k v_{jk}+x_j\right)\delta t_k,\quad v_{jk}=\frac{\delta x_j}{\delta t_k};$$

since the expression at right hand side reads

$$\frac{\left(t_k v_{jk}\right)^2-x_j^2}{t_k v_{jk}-x_j}\delta t_k=\frac{\delta p_{jk}}{\delta\varepsilon_k}\delta\left(\ell^2\right)=\frac{\delta\left(\ell^2\right)}{v_{jk}},\quad \delta p_{jk}=\frac{\hbar}{\delta x_{jk}}$$

$$\delta\left(\ell^2\right)=\left(t_k v_{jk}\right)^2-x_j^2,\quad \delta x_{jk}=t_k v_{jk}-x_j,$$

with notation of (6.4), then

$$\delta\left(x_j t_k\right) = \frac{\delta\left(\ell^2\right)}{v_{jk}} = \frac{v_{jk}^2\,\delta\left(\ell\right)^2}{v_{jk}^3} = \frac{D_{jk}^2}{v_{jk}^3}, \quad D_{jk} = v_{jk}\sqrt{\delta\left(\ell^2\right)}.$$

It is possible to identify here a minimum value of $\delta\left(x_j t_k\right)$ defined by $D_{jk} \to D_{\min}$ and $v_{jk} \to c$, at least asymptotically. Anyway one finds

$$\delta\left(x_j t_k\right)_{\min} = \frac{D_{\min}^2}{c^3}, \quad \delta\left(\ell\right)^2 \equiv \delta\left(\ell^2\right)_{inv}, \quad \delta\left(\ell^2\right)_{inv} = \left(t_k c\right)^2 - x_j^2, \quad v_{jk} = c. \quad (5.26)$$

Whatever the specific value of $D_{\min}$ might be, is interesting the conceptual idea of "granular" space time determined by a minimum linear size of cells, in the present one-dimensional model $\delta\left(length \times time\right)$, that define any macroscopic values of *xt* within these cell; the third (5.26) corresponds to the definition of invariant interval, which is known to be the basis of the special relativity [11].

Accordingly, the Lorentz transformations, in particular, should actually be nothing else but the straightforward consequence of the granular nature of space time.

5.6. Further Comments about the Zero Point Energy

This section generalizes the idea of regarding the zero point energy and volume (5.23) as intrinsic properties of matter, rather than as operative thermodynamic parameters related to specific experimental conditions. According to (5.23), think the zero point volume V_{zp} considering for example an atom surrounded by neighbor lattice atoms; V_{zp} corresponds to its free lattice volume, whatever it might be depending on temperature and mobility. The fact that ϵ_{zp} is defined by the confinement lattice volume around a given atom/ion, implies the limit $\epsilon_{zp} \to 0$ simply because for an isolated free particle $V_{zp} \to \infty$. Nonetheless ϵ_{zp} depends itself on *T*, both because of the thermal dilation of matter that modifies the size of lattice spacing and because the *T* dependent mobility allows one lattice atom to spread well beyond its volume at $T_{\min}$ via the so called "self-diffusion" [12]. Moreover V_{zp} also depends on the presence of lattice defects, which affect the free space available for its delocalization. In particular, one expects that the lattice atom is quenched in one lattice site at $T = T_{\min}$ only; in other words the volume *V* is a thermodynamic parameter experimentally set, whereas V_{zp} is determined by the physics of matter. In the absence of external fields at $T = T_{\min}$, therefore, $kT_{\min}$ is the minimum non-thermal energy of the lattice atom/ion, as it results from all possible interactions with lattice neighbors that determine the available free confinement volume; at $T > T_{\min}$, the thermal energy kT represents actually the additional contribution to the non-thermal zero point energy $\epsilon_{zp} = \epsilon_{zp}\left(T\right)$, so that it seems reasonable to think that in general the simple kT should be implemented as $kT + \epsilon_{zp}$. This holds in particular for the FD and BE statistical distributions. In the case of a single free particle this does not hold, as its delocalization volume is infinity by definition: but in general, when considering the thermodynamic properties of a body of matter, ϵ_{zp} cannot be longer omitted at least in principle.

To justify the legitimacy of this conclusion, consider first an ideal gas inside which energy exchanges occur via direct collisions between its molecules only. Without hypothesizing specific interactions between molecules, e.g. long range Coulomb or dipole interactions between electron shells, holds between p and ϵ of each molecule the general equation

$$\epsilon = \frac{pc^2}{v} = \frac{hc^2\tau}{\lambda\delta x}; \tag{5.27}$$

the second equality introduces the time lapse τ between any successive shocks, during which the molecule travels freely the distance $v\tau = \delta x$. During each τ therefore ϵ remains constant, since inside the gas the energy changes are supposedly due to direct collisions only. In particular, as concerns the zero point energy,

$$\mathcal{E}_{zp} = \frac{hc^2\tau}{\mathcal{V}^{2/3}}, \quad \mathcal{V} = (\lambda\delta x)^{3/2};$$

the notation is justified by the free volume $\mathcal{V}$ dependence analogous to that of the Equation (P77). The fact that $\mathcal{V}$ involves δx is not surprising; δx^{-1} comes from ϵ in (5.27) and agrees with $\delta p/\hbar$, being δp the total range of momentum change $p-(-p)$ after one shock between molecules.

Let ℓ be the distance between one molecule just after its last shock and the wall of the recipient containing the gas; in general $\ell \neq \delta x$. Moreover, as $-\partial\varepsilon/\partial\lambda =$ *force*, (5.27) yields

$$\mathcal{F} = \frac{\epsilon}{\lambda} = \frac{hc^2\tau}{\mathcal{V}}, \quad \mathcal{V} = \lambda^2\delta x, \tag{5.28}$$

where $\mathcal{F}$ is the impact force of the concerned molecule against the wall. If A is the surface of the wall, then the pressure due to the shock of one molecule is

$$\mathcal{P} = \frac{\mathcal{F}}{A} = \frac{hc^2\tau}{A\mathcal{V}}. \tag{5.29}$$

If ℓ is such that $\ell A = \mathcal{V}_\ell$, then the former equation becomes $\mathcal{P} = hc^2\tau\ell/\mathcal{V}_\ell\mathcal{V}$; so, since the numerator has physical dimensions *energy*×*volume*, the result reads

$$\mathcal{P} = \frac{\epsilon_o\mathcal{V}_o}{\mathcal{V}_\ell\mathcal{V}}, \quad \epsilon_o\mathcal{V}_o = hc^2\tau\ell. \tag{5.30}$$

Note that in general neither τ nor ℓ are necessarily constants independent of time; so the pressure $\mathcal{P}$ inferred in (5.30) could be variable during subsequent time lapses τ. Also note that to infer this result in the present one dimensional model it is enough to think the plane A orthogonal to the space coordinate δx. In a general three dimensional approach one should integrate over all possible incidence angles of the molecule against the wall to obtain the pressure, as it is well known. This would entail a numerical factor, which however can be included in $\epsilon_o\mathcal{V}_o$ and thus is irrelevant for the present purposes. Moreover, it is still possible to define a statistical value of $\mathcal{P}$ averaging the shocks over several time lapses during a time range $\Delta t \gg \tau$. So is defined the quantum pressure

$$\mathcal{P}_{zp} = \frac{\epsilon_o \mathcal{V}_o}{\mathcal{V}_{zp}^2}, \quad \mathcal{V}_{zp}^2 = \overline{\mathcal{V}_\ell \mathcal{V}}, \quad \mathcal{V}_{zp} = \mathcal{V}_{zp}(T).$$

This result has been obtained considering the delocalization volume of one molecule; it holds in general for any number N of molecules regarding $\mathcal{V}_{zp}$ no longer as volume of a single molecule but as total experimental volume V_{exp} of ideal gas simply with the position $\mathcal{V}_{zp} = V_{exp}/N$. The right hand side is the average confinement size of the molecule. It is significant to conclude that the pressure of the gas must be expressed not only taking into account the variable dynamical parameter P_{exp} experimentally determined but adding to this latter the contribution $\mathcal{P}_{zp}$ having merely quantum nature; in effect the present reasoning waives considerations about any state equation of ideal or real gases. Also, regarding the macroscopic volume V_{exp} as a further dynamical parameter experimentally set, it contains the quantum contribution $N\mathcal{V}_{zp}$. So, extending the reasoning carried out for one gas molecule to the case of N molecules, one finds

$$P_{eff} = P_{exp} + \frac{fN^2}{\mathcal{V}_{zp}^2}, \quad V_{eff} = V_{exp} - N\mathcal{V}_{exp}, \quad f = \epsilon_o \mathcal{V}_o. \tag{5.31}$$

The volume $\mathcal{V}_{zp}$ is easily understandable, being intuitively evident that the molecules have finite size contributing to the total volume V_{exp} experimentally measurable. Even $\mathcal{P}_{zp}$ is guessable: if the zero point energy is simulated by an oscillator characterized by a non-thermal vibrational frequency that determines the zero point energy, see the next Equation (5.22), then the energy of any oscillator in the gas volume defines a confinement non-thermal energy density equivalent to pressure, see next Equations (8.34). Indeed (5.31) shows that even at $P_{exp} = 0$, e.g. the core of a free body of matter in the vacuum with zero applied pressure, there is a residual internal pressure, non-eliminable, e.g. it could act substantially similarly to the repulsion between electron shells of molecules; moreover the latter equation shows even a non-reducible residual volume $\mathcal{V} = N\mathcal{V}_o$. The relationship between energy, pressure and volume follows directly from (5.29) as $\mathcal{P}\mathcal{V} = \hbar c^2 \tau / A = const$ at fixed temperature: indeed at the right hand side of the first equality appear fixed quantities, of course at given ℓ and τ and thus constant T. So

$$P_{eff} V_{eff} = \left(\mathcal{P}_{zp} + \frac{fN^2}{\mathcal{V}^2} \right) \left(\mathcal{V} - N\mathcal{V}_{zp} \right).$$

This equation reminds closely the characteristic terms of the Van der Waals equation, where f and $\mathcal{V}_{zp}$ are approximately regarded as gas constants; this holds also here, even though the pressure and volume terms (5.31) have been inferred considering initially an ideal gas via quantum considerations about its constituting molecules. The interactions between these molecules, even not hypothesized and purposely introduced "a priori", appear as quantum effects regardless of specific considerations about their actual nature, which is in effect "hidden" in parameters like τ or ℓ or f descriptive of the properties of the gas; hence is not surprising that the coefficients $\epsilon_o\mathcal{V}_o$ and $\mathcal{V}_o$ contribute to that characterizing the famous Van der Waals equation. The conceptual reasons underlying this equation are well known; is interesting however that the form of the resulting equation based on the present approach is analogous to that just found.

These considerations are now extended to the concept of temperature once having introduced the quantum meaning of $T_{\min}$ related to ϵ_{zp} to show that the zero point energy, typical quantum effect, affects the macroscopic properties of gases: ideal gas is the one where these quantum effects are approximately neglected along with the long range mutual interactions as well.

If effectively exists a minimum temperature $\neq 0$, then it must be defined by $kT_{\min} = \epsilon_{zp}$. However, the right hand side is in general function of T itself; indeed it has been shown that $\epsilon_{zp} \propto V_l^{-2/3}$, being V_l the lattice volume available around a given lattice site. So the thermal dilation modifies ε_{zp}, whose non-thermal physical meaning however still holds identically. Moreover the local mobility is itself T dependent, as it is intuitive to think; in effect it is known that by self diffusion, atoms in a given lattice point can exchange of place with lattice neighbors, so that the actual volume allowed to a given atom is increased by the number of neighbor elementary cells accessible.

These considerations should be also extended in particular to the statistical distributions of bosons and fermions, usually written as a function of kT only: taking into account the considerations elucidated in the case of the Van der Waals equation, one should conclude that strictly speaking in the case of a solid body the simple term kT should be replaced by $k\left(T+T_q\right)$, where kT_q accounts for the quantum contribution related to the T-dependent zero point energy with $T_q = T_q\left(T\right)$.

$$\frac{w}{q_{\pm}} = \frac{W_0}{\exp\left(\left(\mu-\mu_0\right)/k\mathcal{T}\right)\pm 1}, \quad \mathcal{T} = nT_{\min} + \epsilon_{zp}/k \qquad (5.32)$$
$$\epsilon_{zp} = \epsilon_{zp}\left(T\right), \quad n = integer \geq 1.$$

Even though the value of $T_{\min}$ is presumably much lower than the ordinary temperatures today attainable and experimentally measurable, the Equation (5.25) suggests the chance of being tested in a situation where ϵ_{zp} is relevant, *i.e.* in the case of theoretical models of solid state physics. In effect the paper [13] implements the ideas of quantized temperature and statistical distributions (5.32), both introduced as hypotheses; the specific heat calculated agrees very well with the experimental data from very low T up to the melting point for several metals with different crystal structures.

6. Some Relativistic Corollaries

In this section are examined some relevant relativistic corollaries of the previous results. The importance of the following considerations is shown by the chance of obtaining contextually relativistic results in the same conceptual frame of the quantum results previously obtained. The next two subsections emphasize the importance of the previous equations (3.15) and (4.9), now again under test after their previous validations, see respectively Equations (3.20) and (4.6), (4.7), (4.13) to (4.18).

6.1. The Invariant Interval

Implement the Equation (3.15), once more under test after the result (3.20), rewritten as follows

$$\Delta\epsilon = c\Delta\left(pr\right), \quad \Delta\epsilon = \sigma_\epsilon \varepsilon - \frac{b\varepsilon^2}{c^2} - \epsilon_a, \quad r = \sigma_p - ap/c \tag{6.1}$$

depending on whether $r \lesseqgtr 1$ one has $\Delta\epsilon \gtreqless c\Delta p$. Hence, squaring both sides, $\left(\Delta\epsilon\right)^2 \gtreqless \left(c\Delta p\right)^2$ reads

$$\left(c\Delta p\right)^2 - \left(\Delta\epsilon\right)^2 = K, \quad K \gtreqless 0, \tag{6.2}$$

where anyway K is the resulting value from the left hand side of the Equation (6.2).

Consider first $K > 0$.

Implementing (3.1), one finds

$$\hbar^2 c^2 \left(\Delta x\right)^{-2} - \hbar^2 \left(\Delta t\right)^{-2} = K,$$

whence

$$\frac{\hbar^2}{\left(\Delta t\Delta x\right)^2}\left(c^2\Delta t^2 - \Delta x^2\right) = K; \tag{6.3}$$

this equation reads identically

$$\Delta\left(s^2\right) = \frac{K}{\hbar^2}\left(\Delta x\Delta t\right)^2 = \left(c\Delta t\right)^2\left(1 - \frac{v^2}{c^2}\right), \quad \Delta\left(s^2\right) = s'^2 - s''^2 = c^2\Delta t^2 - \Delta x^2, \tag{6.4}$$

having implemented once more $v = \Delta x/\Delta t$, whereas

$$\Delta x^2 = \frac{\left(\hbar c\right)^2}{K}\left(1 - \frac{v^2}{c^2}\right)$$

reads by dimensional reasons, whatever the value of K might be,

$$\Delta x'^2 = \frac{\Delta x^2}{1 - v^2/c^2}, \quad \Delta x'^2 = \frac{\left(\hbar c\right)^2}{K}. \tag{6.5}$$

It is easy to recognize the Lorentz transformation of the intervals $\Delta x'$ and Δx in two different inertial reference systems R and R', hence in (6.4) both $\Delta\left(s^2\right)$ and $\Delta x\Delta t$ must be invariants, as found in particular in (3.62); indeed K is invariant itself if r of (6.1) is calculated via invariant forms of ε and p, see Equations (3.64).

Consider now $K < 0$.

The interval defined in (6.3) reads $\Delta x^2 - c^2\Delta t^2 > 0$; moreover (6.2) reads $\Delta\epsilon^2 = (c\Delta p)^2 - K$, whence (3.20) with an appropriate value of *K*. Also this result holds regardless of the local limit condition $\Delta \to \partial$; so this is not a local property, but a feature of the whole space time. It is significant the fact that these results have been obtained with the help of the quantum uncertainty relationships (3.1).

6.2. The Gravity Force

Consider again the Equation (4.9), now once more under test after its early validation via the Equations (4.6), (4.7) and (4.13) to (4.18). Implementing the Equations (3.1) and the initial definition (2.2) of v one finds

$$\Delta t = \frac{v^2}{c^2}\frac{\delta x \delta p}{\delta \varepsilon} = \frac{\delta x^3}{c^2 \delta t^2}\frac{\delta p}{\delta \varepsilon}$$

that yields, after multiplying and dividing the right hand side by an arbitrary mass *m*,

$$mc^2 = \mathcal{G}\frac{m}{\ell}, \quad \frac{1}{\ell} = \frac{\delta p/\Delta t}{\delta \varepsilon} = \frac{F}{\delta \varepsilon}, \quad \mathcal{G} = \frac{\delta x^3}{m\delta t^2}; \tag{6.6}$$

by dimensional reasons ℓ is an arbitrary length. The function $\mathcal{G}$ consists in general of a constant term G plus a correction term $\mathcal{R}$; indeed $\mathcal{G}$ can be expanded in series around arbitrary reference ranges δx_0, δt_0 and m_0 defining G, *i.e.*

$$\mathcal{G} = G + \mathcal{R}, \quad G = \frac{\delta x_0^3}{m_0 \delta t_0^2}, \quad \mathcal{R} = \sum_{j=1}^{\infty} a_j u^j, \quad u = \frac{\delta x^3}{m\delta t^2} - \frac{\delta x_0^3}{m_0 \delta t_0^2}, \tag{6.7}$$

being a_j appropriate coefficients of the power series expansion and $\mathcal{R}$ the sum of the higher order terms of the series. This means calculating $\mathcal{G}$ around an arbitrary value $\mathcal{G}_0 \equiv G$. So (6.6) reads

$$c^2 = (G + \mathcal{R})\frac{m}{\ell}, \tag{6.8}$$

It is evident that the general relativity appears in this result: the Newtonian potential Gm/ℓ recognizable in this equation is the approximate particular case of a more general equation involving $\mathcal{R}$ too. It is known for example that the simple addition of a further term to the Newtonian potential is enough to calculate correctly the perihelion precession even without implementing the basic assumptions and tensor calculus of general relativity [14]. Unfortunately the Newton physics does not justify itself this additional term. Nonetheless the mere series expansion of the last (6.6) around an arbitrary space time constant term G legitimates the chance of generalization without introducing “ad hoc” hypotheses. In this respect some further considerations exposed below regard in particular the additional non-Newtonian terms due to the series (6.7), still in the conceptual frame introduced by (1.11) and (1.12) as done throughout this paper. In effect, some papers among which [15] show that the quantum uncertainty allows to infer the most relevant results of general relativity in a unique frame that includes of course the quantum physics. The presence of c^2 at the left hand side of (6.8) suggests the chance of multiplying both sides by a further arbitrary mass $\pm m'$; one finds then

$$\mp\mathcal{R}\frac{m'm}{\ell}\pm m'c^2=\epsilon_G=\pm G\frac{m'm}{\ell}, \qquad (6.9)$$

which defines two important quantities

$$\phi_G=\frac{\epsilon_G}{m'}=\pm\frac{mG}{\ell}, \quad F_G=-\frac{\delta\epsilon_G}{\delta\ell}\approx\pm G\frac{m'm}{\ell^2}, \quad \mathcal{R}\ll G. \qquad (6.10)$$

At the left hand side of (6.9) appears the rest energy $\pm m'c^2$ of m' plus a correction term resulting from the series expansion of *G*, at the right hand side the related potential energy ϵ_G to which correspond the pproximate gravitational potential ϕ_G and force F_G of (6.10) via the constant zero order term $\mathcal{G}$ only. Note that the Equations (2.35) and (3.64) have introduced the concept of states of negative energy; hence the left hand sides of (6.9) and (6.10) have physical meaning even regarding $\pm m'$ as unique mass $m_\pm$ with positive and negative energy states $\pm m'c^2=m_\pm c^2$. So the Equations (4.9) and (3.1) prospect a possible chance for positive sign of m', *i.e.* a repulsive gravity force as already found in [16] [17] [18], whereas (6.9) yields

$$\phi_G^*=\frac{mG}{\ell}, \quad F_G^*=-\frac{\delta\epsilon_G^*}{\delta\ell}\approx G\frac{m'm}{\ell^2}. \qquad (6.11)$$

Consider here the negative value $-m'$ and suppose $\mathcal{R}\ll G$; (6.9) and (6.10) yield

$$\phi_G=-\frac{mG}{\ell}, \quad F_G\approx -G\frac{m'm}{\ell^2}. \qquad (6.12)$$

The Equations (6.12) and (6.11) will be further explained just below, see the next Equation (8.6), after having first validated the results achieved in this subsection.

1) The Equation (6.8) reads also

$$\ell=\frac{G+\mathcal{R}}{G}\frac{mG}{c^2},$$

which in the particular case $\mathcal{R}=G$ implies a specific value of ℓ given by

$$\ell_{bh}=\frac{2mG}{c^2}. \qquad (6.13)$$

To understand the physical meaning of this result, rewrite identically (6.8) multiplying both sides by an arbitrary factor $\sigma\le 1$; then

$$v_e^2=\mathcal{R}_o Gm/\ell, \quad \mathcal{R}_o=(1+\mathcal{R}/G)\sigma, \quad v_e^2=\sigma c^2.$$

For $R_o=2$, in particular, one recognizes the well known escape velocity v_e of an arbitrary mass m' at a distance ℓ from the gravitational center of mass of *m*, also inferable via (6.12) under the condition of null total energy (potential plus kinetic) of m' at infinity. As $R_o=2$ is compatible with $\sigma=1$ simply putting $\mathcal{R} = G$ in (6.8), this limit condition for the chance of m' at distance ℓ of escaping from the gravity field of m holds also for c. This velocity can be nothing else but that of a photon, so (6.13) shows that at any $\ell' \le \ell_{bh}$ even light cannot escape from the gravity field of m. Trivial manipulations of the early Equations (4.8) and (4.9) yield therefore the black hole limit condition between mass m and distance ℓ.

2) The present way to introduce the gravity force explains why any test mass m' behaves in the same way in the field of a source mass m: the masses m of (6.6) and m' of (6.9) have been introduced subsequently and independently each other. Note the conceptual difference between the present reasoning and that of the Newtonian approach: Newton has contextually introduced two masses to define their mutual interaction law, here the masses have been consequentially introduced starting from (6.6) because c^2 requires the concept of mass to introduce that of gravitational energy. It is not surprising that once having decided either mass as a field source, the behavior of the other is uniquely fixed: as both masses are independent and arbitrary, once having fixed m the behavior of m' is uniquely determined. In other words there is no reason to expect that any other mass m'' behaves in a different way from m' in the gravitational field of m because the law governing its dynamics has already been independently fixed.

3) Here m and m' are regarded in fact as gravitational and inertial masses; the first one defines the gravitational potential ϕ_G, the second one defines their mutual interaction force F_G. As they are interchangeable, their gravitational and inertial role is physically equivalent and indistinguishable. Thus gravitational and inertial masses must be equal.

4) Consider that (6.9) and (6.12) yield

$$\frac{\ell_{bh}}{\ell} = -2\frac{\phi_G}{c^2}, \quad \delta\left(\frac{\ell_{bh}}{\ell}\right) = -\frac{\ell_{bh}}{\ell^2}\delta\ell = -2\frac{\delta\phi_G}{c^2}. \tag{6.14}$$

Identify now the distance ℓ with one wavelength of a light beam propagating in the vacuum; thus

$$\ell = \frac{c}{\nu}, \quad \delta\ell = -\frac{c}{\nu^2}\delta\nu, \quad \frac{\delta\ell}{\ell^2} = -\frac{c\delta\nu}{\nu^2}\frac{\nu^2}{c^2} = -\frac{\delta\nu}{c}.$$

Hence, given a light wave propagating at distance $\ell > \ell_{bh}$, one finds

$$-\frac{\ell_{bh}}{\ell^2}\delta\ell = -2\frac{\delta\phi_G}{c^2} = -\ell_{bh}\frac{\delta\nu}{c};$$

then

$$\frac{\delta\phi_G}{c^2} = \frac{\delta\nu}{\nu_r}, \quad \nu_r = \frac{2c}{\ell_{bh}} = \frac{c^3}{mG}, \quad \delta\nu = \nu - \nu_r, \tag{6.15}$$

being ν_r a reference frequency of the wave. The first equation defines the famous red shift $(\nu - \nu_r)/\nu_r$ of a light wave due to the gravitational potential field change $\delta\phi_G$.

5) It is instructive to consider now two photons freely moving in the vacuum

on the diametric plane of two concentric hyperspheres: the inner photon just at radial distance ℓ_{bh} from the gravity center of *m*, given by (6.13), the outer photon at any radial distance $\ell > \ell_{bh}$ from *m*. The previous result shows that the inner photon cannot escape from the gravity field of *m*, so it can move on the surface of the inner hypersphere only, whereas any photon moving at very large distance $\ell \gg \ell_{bh}$ from *m* is free to travel unperturbed as a limit case from minus infinity to infinity or vice-versa. Is reasonable the idea that the outer photon moving at the closest approach distance $\ell > \ell_{bh}$ should follow an intermediate behavior, *i.e.* a curved space time path bent by *m*. This preliminary consideration justifies why the problem of light beam bending is tackled here with reference to the previous Equation (6.13). The standard way to approach the problem considers the curved trajectory traveled by the photon that follows the space time curvature along an arc δs around *m* at distance ℓ; the position of the photon before and after its closest approach to *m* defines the characteristic deflection angle $\varphi = \varphi(\ell, m)$ equal to that formed by the tangents to the osculating circumference at boundaries of δs. From a quantum point of view, however, the concepts of position and trajectory are missing, rather the approach must be similar to that followed to infer (4.16). Just for this reason the present reasoning is instructive to highlight how the quantum requirements are plugged into and provide information on this typical relativistic context.

Regard the arc δs of osculating circumference of radius ℓ defined by $\delta s = \ell\varphi$ conceptually according to of (3.1), *i.e.* as an uncertainty range where the photon is delocalized. Accordingly δs is actually given by two half angles $\delta s_{cw} = \ell\varphi/2$ and $\delta s_{ccw} = -\ell\varphi/2$ traveled by the photon along clockwise and counterclockwise paths around the middle point $\delta s/2$; indeed the photon displacements implied by δs_{cw} and δs_{ccw} are physically indistinguishable, because nothing is known about the motion features of any particle within an uncertainty range. This point of view skips the idea of a photon entering in δs through one boundary and exiting from the other boundary, which in fact would define δs as an element of trajectory. So $\delta s = \delta s_{cw} - (-\delta s_{ccw}) = \ell\varphi$ waives the whole φ, which would imply discriminating the events where the photon travels through δs coming from $-\infty$ towards ∞ or from ∞ towards $-\infty$; actually these events are indistinguishable likewise the boundaries of δs themselves. So, with respect to gravity center of *m*, the angle of physical interest is $\varphi/2$ and not φ to account for the sought total δs. In other words the Equations (3.1) compel merging two half-paths into a unique travel path without discriminating either of them.

Consider then an angle $\varphi/2$ on a circumference of radius ℓ and its related length $(\delta s/2)/\ell$ to describe the uncertainty range δs where is delocalized the photon. Rewrite the second (6.14) as follows

$$\frac{\ell_{bh}}{\ell} - \frac{\ell_{bh}}{\ell_o} = -\frac{\delta s}{\ell} = -2\left(\frac{mG/c^2}{\ell} - \frac{mG/c^2}{\ell_o}\right), \quad -\frac{\ell_{bh}}{\ell}\delta\ell = \delta s;$$

here the uncertainty ranges $\delta(\ell_{bh}/\ell)$ and $\delta(\phi_G/c^2)$ have been simply rewritten with the usual notation of any $\delta y = y - y_o$ by definition, whatever the concerned y might actually be. In this specific case ℓ and ℓ_o are two different distances of the photon from the gravity center of m. It is clear that it is convenient to put here $\ell_o \to \infty$ because, as previously stated, we are interested to describe the situations where one photon initially unperturbed passes at a finite distance ℓ from m. Hence the last equation reads

$$\frac{\ell_{bh}}{\ell} = -\frac{\delta s}{\ell} = -2\frac{mG/c^2}{\ell}.$$

If, for the aforesaid quantum reasons $\delta s/\ell = \varphi/2$ on the osculating circle, one finds immediately

$$\frac{\varphi}{2} = \frac{l_{bh}}{\ell} = 2\frac{mG}{c^2\ell}.$$

These simple considerations emphasize the actual quantum character of one of the most representative relativistic predictions, the gravitational lensing; the famous factor 4 defining φ at the right hand side appears to be actually the fingerprint of the quantum uncertainty.

6) Consider the physical dimensions of the gravity constant inferred from (6.6): according to (6.7), space and time range sizes δx_0 and δt_0 concur to its macroscopic value together with the arbitrary mass m_0. From a quantum point of view, (6.7) does not exclude the chance of mass fluctuation, *i.e.* according to (4.6)

$$\begin{aligned} \delta G &= -\frac{\delta x_0^3}{m_0^2 \delta t_0^2}\delta m_0 + \frac{1}{m_0}\delta\left(\frac{\delta x_0^3}{\delta t_0^2}\right), \\ \delta m_0 &= m_0' - m_0 = m_0\left(1 - \sqrt{1 - \frac{v^2}{c^2}}\right). \end{aligned} \tag{6.16}$$

Even considering preliminarily the fluctuation of m_0 only, and thus energy fluctuation $\delta\epsilon^* = c^2\delta m_0$ only, the constancy of G is expressible as

$$\delta G = 0, \quad \frac{\delta x_0^3}{m_0 \delta t_0^2}\delta m_0 = m_0\delta\left(\frac{\delta x_0^3}{\delta t_0^2}\right)$$

i.e. $\delta\epsilon^*$ implies the change of $\delta x_0^3/\delta t_0^2$ as well. Note that δx_0 and δt_0 are not usual dynamical variables characterizing physical laws, rather they define the structure of G itself. In effect, this quantum standpoint implies that transient fluctuations of values of G are in principle possible, being compatible with corresponding space time energy quantum fluctuations $\delta\epsilon^*$. Making explicit the right hand side of this last equation, trivial algebraic steps yield

$$\begin{aligned} \delta G &= 3\delta x_0 \frac{x_0^2}{m_0 t_0^2} - 2\delta t_0 \frac{x_0^3}{m_0 t_0^3} - \delta m_0 \frac{x_0^3}{m_0^2 t_0^2} \\ &= \frac{x_0^2 \delta x_0}{m_0 t_0^2}\left(3 - 2\frac{x_0}{t_0}\frac{\delta t_0}{\delta x_0} - \frac{x_0}{m_0}\frac{\delta m}{\delta x_0}\right) = 0 \end{aligned}$$

so that $G = const1$ simply if

$$\frac{x_0}{t_0}=\sigma_1\frac{\delta x_0}{\delta t_0},\quad \frac{m_0}{x_0}=\sigma_2\frac{\delta m_0}{\delta x_0},\quad \frac{m_0}{t_0}=\frac{\delta m_0}{\delta t_0},\quad 2\sigma_1+\frac{1}{\sigma_2}=3;$$

otherwise, e.g. a different correspondence between σ_1 and σ_2 whatever their values might be, the quantum definition of G admits $\delta G\neq 0$. All this has clearly to do with the existence of inflationary era of the early Universe: an appropriate fluctuation $\delta G\neq 0$ of G can contribute in principle to the sudden increase of expansion rate of the early Universe. Among the implications of these assertions, one deserves particular attention: the possible fluctuations of G affect the black hole length ℓ_{bh} of (6.13) of a given m. Unfortunately further discussion on this crucial point is outside the aims of the present paper.

7) Returning now to the Equations (5.25), the maximum value of momentum wavelength λ_{zp} has been related to a suitable space range Δr defined as the radius of a hypersphere within which is delocalized an arbitrary amount of mass. As the steady wavelength appropriate to calculate the zero point energy $\epsilon_{zp}^{\min}$ defining $T_{\min}$ corresponds to the maximum delocalization extent physically conceivable in our Universe, is reasonable to relate λ_{zp} to the diametric size $2\Delta r_u$ of the Universe, regarded here as a hypersphere with diameter $2\Delta r_u$. So think the mass ideally bouncing within one diametric distance, whose maximum space extent corresponds to one half wavelength; this wave has thus steady nodes at the diametric boundaries of the Universe, regarded in effect as a hypersphere. Replacing Δr of the Equation (5.25) with Δr_u and putting $v\to c$ asymptotically, one finds

$$\epsilon_{zp}^{\min}=\frac{hc}{4\Delta r_u},\quad T_{\min}=\frac{\epsilon_{zp}^{\min}}{k}=\frac{hc}{4k\Delta r_u},\quad \omega_u=\frac{2\epsilon_{zp}^{\min}}{\hbar}=\frac{\pi c}{\Delta r_u}. \tag{6.17}$$

The current estimate $\Delta r_u=4.35\times 10^{28}$ cm [19] yields the numerical values

$$\epsilon_{zp}^{\min}=1.2\times 10^{-45}\ \text{erg},\quad T_{\min}=8.3\times 10^{-30}\ \text{K},\quad \omega_u=2.2\times 10^{-18}\ \text{s}^{-1} \tag{6.18}$$

In addition to the preliminary remarks about the Equation (5.25) previously emphasized, these numerical results suggest further implications:

- Finite Universe means identically $\epsilon_{zp}^{\min}\neq 0$.
- The limit $T_{\min}\to 0$ would hold in an infinite universe only.
- To guess the physical meaning of the small value of ϵ_{zp}, note that $1/\omega=4.6\times 10^{17}$ s fits surprisingly well the estimated age of the universe 4.35×10^{17} s reported in [19].
- The fact that the energy corresponding to λ_{zp} agrees reasonably with the estimated order of magnitude of the age of the universe, suggests that $\epsilon_{zp}^{\min}$ with which has been calculated $T_{\min}$ could be a possible vacuum energy fluctuation, still in progress, of the whole Universe.
- It is interesting that ω_u that determines $T_{\min}$ agrees surprisingly well with the Hubble constant.

8) It is possible to implement these results to calculate another important property of the Universe, *i.e.* the vacuum energy density η_u. In general, the energy density corresponds from the dimensional point of view to

$mass \times frequency^2 / length$. Consider now that according to (6.13) nothing, even the light, can escape from a range size ℓ_{bh} enclosing the mass m; hence, the Equation (6.13) represents a significant opportunity to describe how to trap inside a volume of space time energy that cannot be irradiated nor lost outside it. After having inferred that an energy $\hbar\omega_u$ pervades all Universe, whose zero point value $\hbar\omega_u/2$ determines the zero point energy defining $T_{\min}$, is attracting the idea that η_u can be calculated just with the value of ω_u controlling $T_{\min}$. This idea links the vacuum energy density to the zero point energy (5.22) related to $kT_{\min}$. If so, then calculating $m/\ell = c^2/2G$ according to (6.13) and implementing ω_u just calculated,

$$\eta_u = \frac{c^2}{2G}\omega_u^2 = 3.3\times10^{-8}\ \text{erg}/\text{cm}^3 = 3.3\times10^{-9}\ \text{J}/\text{m}^3. \qquad (6.19)$$

The sensible value of vacuum energy density further supports the way to calculate the values (6.18). This means that the concept of vacuum does not imply that of "nothing": rather the vacuum must consist of virtual particles whose energy, *i.e.* mass, governs the residual vacuum energy density in agreement with the third law of thermodynamics.

9) But there is more. Implement the mass M_u of the part of visible Universe [16] estimated counting the stars only to calculate

$$M_u = 3\times10^{55}\ \text{g}, \quad \frac{2M_uG}{c^2} = 4.5\times10^{27}\ \text{cm}. \qquad (6.20)$$

On the one hand one expects that M_u is estimated by defect, without taking into account that other forms of energy distributed in the Universe could in principle increase this value; whatever this additional energy ϵ^* might be, it concurs with M_u by ϵ^*/c^2 to the total mass of the Universe. On the other hand are visible only the stars whose distance does not overcome the observation limit posed by the light speed; assuming that the distribution of galaxies and thus stars is uniform in the Universe, the actual mass due the total number of stars should be

$$M^* \approx 3\times10^{55}\left(\frac{\Delta r_u}{c\Delta t_u}\right)^3 = 1.1\times10^{57}\ \text{g}.$$

This value still estimates the total visible matter of the Universe as if the light speed would be infinite; as such, however, it does not tell anything about other possible contributions inherently "dark", *i.e.* non-luminous, for example the vacuum energy/c^2 or the zero point energy/c^2. Compare thus just this value M^* with the total mass related to the whole vacuum energy density η_u calculated above. It is

$$\frac{4}{3}\pi\Delta r_u^3\frac{\eta_u}{c^2} = 1.3\times10^{58}\ \text{g}$$

The total vacuum energy of the Universe is still about 12 times higher than that of all visible objects. Make at this point a hypothesis:

The energy density of vacuum and that of matter are equal, *i.e.* regard matter and vacuum as two different thermodynamic systems at global equilibrium.

So M^* must be incremented just by this factor to make equal the respective densities. Write thus the total mass balance as a function of the true visible mass as

$$2M^* + M_p + M_\varepsilon = M_{tot} = 12M^* \tag{6.21}$$

the factor 2 accounts for the antimatter, wherever it might be in the Universe, whereas M_p and M_ε are the missing masses also concurring to the factor 12 assumed true. The notations account for the fact that the concept of mass can be defined in general via p/v and ε/v^2. Rewrite now these positions with the help of (2.36)

$$M_p = \frac{h}{\lambda_p v}, \quad M_\varepsilon = \frac{h\nu_\varepsilon}{v^2}, \quad \nu_\varepsilon = \frac{v}{\lambda_\varepsilon} \tag{6.22}$$

where λ_p and ν_ε are the pertinent momentum wavelength and energy frequency. Thus

$$M_p = \frac{v_o}{v}\frac{h}{\lambda_p v_o}, \quad M_\varepsilon = \left(\frac{v_o}{v}\right)^2 \frac{hv}{\lambda_\varepsilon v_o^2}. \tag{6.23}$$

being v_o arbitrary velocity. Moreover rewrite identically (6.21) as

$$\begin{gathered} 2M^* + q_p M^* + q_\varepsilon M^* = 12M^*, \quad M_p = q_p M^*, \\ M_\varepsilon = q_\varepsilon M^*, \quad M^* = \frac{h}{\lambda_p v_o} = \frac{hv}{\lambda_\varepsilon v_o^2}, \end{gathered} \tag{6.24}$$

where q_p and q_ε are appropriate coefficients able to express numerically M_p and M_ε via M^*. Actually the physical meaning of these positions is to establish a relationship between visible mass and the other contributions to M_{tot}. The last position, in particular, is possible because λ_p and λ_ε are arbitrary. One equation to determine these coefficients is obviously

$$q_p + q_\varepsilon = 10, \quad q_p = q_p(x,t), \quad q_\varepsilon = q_\varepsilon(x,t). \tag{6.25}$$

Moreover, as (6.23) reads

$$\frac{v_o}{v} = q_p, \quad \left(\frac{v_o}{v}\right)^2 = q_\varepsilon = q_p^2, \quad \frac{v}{\lambda_\varepsilon} = \frac{v_o}{\lambda_p}, \tag{6.26}$$

(6.24) yields

$$q_p M^* + q_\varepsilon M^* = \left(q_p + q_p^2\right) M^*, \quad q_p + q_p^2 = 10, \tag{6.27}$$

and thus $q_p = 2.7$ and $q_\varepsilon = 7.3$. Being the third position (6.26) certainly fulfilled via the arbitrary v_o, which however does not appear in (6.27), (6.24) and (6.21) become

$$2M^* + 2.7M^* + 7.3M^* = 12M^*, \quad M_p = 2.7M^*, \quad M_\varepsilon = 7.3M^*,$$

which are expressed more significantly in relative $\%$:

$$(8.3 + 8.3)\%M^*, \quad 22.5\%M_p, \quad 60.8\%M_\varepsilon. \tag{6.28}$$

The papers [10] propose a possible explanation, here omitted for brevity, about why matter and antimatter are separated. It is more important to note that there is no "ad hoc" hypothesis in this reasoning, rather a further implication of the fundamental concept of uncertainty repeatedly invoked throughout all this paper and again exploited here through the assumption of vacuum/matter equilibrium. In effect the mass densities calculable via the terms M_p and M_ε additional to M^* in (6.21) correspond to the respective energy densities and thus to pressure terms inside the Universe. The next Equations (8.28) to (8.34) elucidate this point.

10) Multiply (6.6) by the mass m_a introduced in (3.20) and (3.21); recalling the second (3.26) it is possible to write

$$m_a c^2 = \frac{2\epsilon_a}{n'\sigma_\varepsilon^2 + n^{-2}} = G\frac{m_a m}{\ell}, \tag{6.29}$$

which links via c^2 the Newtonian gravity and quantum energy at the left hand side. Whatever σ_ε might be, the arbitrary distance ℓ is repalced by arbitrary quantum numbers n and n'. In this sense (3.20) is reasonably defined in [6] as equation of quantum gravity.

It is true that actually G should be replaced by $\mathcal{G}$ to plug all considerations carried out in the subsection 6.2 into the relativistic realm via (6.7); but it is also true that actually m_a has been defined in order to make σ_ε and σ_p of Equation (3.15) consistent with the harmonic oscillator form (3.25) consequent (3.24). Modifying the definitions (3.23) in order to define non-harmonic oscillations would mean adding additional correction terms corresponding to the higher order terms of the series (6.7). It is possible to say shortly that the Newton gravitational energy corresponds to quantum harmonic oscillators at the left hand side of (3.25), the relativistic gravity is described by non-harmonic oscillator terms replacing the mere $n'h\nu$ of (3.25).

7. Klein Gordon, Proca and Maxwell Equations

Implement now the Equations (3.30) introducing a function of $\psi^{(0)}$ defined as follows

$$\psi^{(0)} = Q\exp(\sigma\psi), \quad \psi^{(0)} = \psi^{(0)}(x',t), \tag{7.1}$$

where σ and Q are arbitrary constants. So the first Equation (3.30) yields

$$\frac{1}{\psi^{(0)}}\frac{\delta^2\psi^{(0)}}{\delta t^2} = \sigma\left(\frac{\delta^2\psi}{\delta t^2} + \sigma\left(\frac{\delta\psi}{\delta t}\right)^2\right) = \sigma(1\pm\sigma)\frac{\delta^2\psi}{\delta t^2}; \tag{7.2}$$

proceeding in analogous way with the help of (3.31), the second Equation (3.30) yields

$$\frac{1}{\psi^{(0)}}\frac{\delta^2\psi^{(0)}}{\delta x'^2} = \sigma(1\pm\sigma)\frac{\delta^2\psi}{\delta x'^2}. \tag{7.3}$$

Hence, replacing (7.3) and (7.2) in the right hand side of (3.32), the result is an equation expressed as a function of the functional $\psi^{(0)}$ of ψ

$$\left(\frac{m'c}{\hbar}\right)^2\psi^{(0)}=-\frac{\delta^2\psi^{(0)}}{\delta t^2}+\frac{\delta^2\psi^{(0)}}{\delta x'^2},\qquad m'=m\sqrt{\sigma(1\pm\sigma)};\tag{7.4}$$

as m appears at the left hand side only of (3.32), it has been included in m' together with the factor $\sigma(1\pm\sigma)$ without loss of generality. Taking the limit $\delta\to\partial$, which by consequence implies $\psi^{(0)}\to\psi_{KG}$ as well, this one dimensional result is actually the well known Klein Gordon equation

$$\left(\frac{m'c}{\hbar}\right)^2\psi_{KG}=-\frac{\partial^2\psi_{KG}}{\partial t^2}+\frac{\delta^2\psi_{KG}}{\partial x'^2}\to\Box\psi_{KG}+\left(\frac{m'c}{\hbar}\right)^2\psi_{KG}=0,$$

$$\Box=\frac{\partial^2}{\partial t^2}-\frac{\partial^2}{\partial x_i^2}.$$

On the one hand it confirms the validity of the positions (3.30) and thus of (3.39) too; on the other hand this result shows that the Klein Gordon equation is inferred from the local functional (7.1) of the space time function (3.39). The latter equation simply rewrites the former according to the usual 4D d'Alabertian.

This result can be further generalized taking advantage that the last equation is actually expressed as a function of m' and not of the initial m. The fact that m' can take two values depending on either sign in (7.3) suggests that in fact two equations are tacitly implied by the unique Equation (7.1); for example one scalar equation and one vector equation should be compatible with (7.3). This is very easily proven showing that the scalar and vector fields of the Proca equations can be combined into one resulting Klein Gordon-like equation. The most straightforward way to demonstrate this assertion starts just from the Proca equations

$$\Box\phi-\frac{\partial}{\partial t}\left(\frac{1}{c^2}\frac{\partial\phi}{\partial t}+\nabla\cdot\boldsymbol{A}\right)=-\left(\frac{mc}{\hbar}\right)^2\phi,$$

$$\Box\boldsymbol{A}+\nabla\left(\frac{1}{c^2}\frac{\partial\phi}{\partial t}+\nabla\cdot\boldsymbol{A}\right)=-\left(\frac{mc}{\hbar}\right)^2\boldsymbol{A},$$

which actually, owing to the definition of the operator $\Box$, read respectively

$$\nabla^2\phi+\frac{\partial(\nabla\cdot\boldsymbol{A})}{\partial t}=\left(\frac{mc}{\hbar}\right)^2\phi,\quad \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2}+\frac{1}{c^2}\frac{\partial(\nabla\phi)}{\partial t}=-\left(\frac{mc}{\hbar}\right)^2\boldsymbol{A};\tag{7.5}$$

thus ϕ and $\boldsymbol{A}$ are the sought scalar and vector fields linking the two equations. Multiply the former equation by an arbitrary function $f=f(x,y,z,t)$ and the latter by an arbitrary velocity vector $\boldsymbol{v}_o=\boldsymbol{v}_o(x,y,z,t)$, *i.e.*

$$f\nabla^2\phi+f\frac{\partial(\nabla\cdot\boldsymbol{A})}{\partial t}=\left(\frac{mc}{\hbar}\right)^2f\phi,$$

$$\boldsymbol{v}_o\cdot\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2}+\boldsymbol{v}_o\cdot\frac{1}{c^2}\frac{\partial(\nabla\phi)}{\partial t}=-\left(\frac{mc}{\hbar}\right)^2\boldsymbol{v}_o\cdot\boldsymbol{A}.$$

Subtracting side by side these equations,

$$f\left(\nabla^2\phi+\frac{\partial(\nabla\cdot\boldsymbol{A})}{\partial t}-\left(\frac{mc}{\hbar}\right)^2\phi\right)=\boldsymbol{v}_o\cdot\left(\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2}+\frac{1}{c^2}\frac{\partial(\nabla\phi)}{\partial t}+\left(\frac{mc}{\hbar}\right)^2\boldsymbol{A}\right),\tag{7.6}$$

trivial algebraic steps yield the following scalar equation

$$f\nabla^2\phi + f\frac{\partial(\nabla\cdot\boldsymbol{A})}{\partial t} - \boldsymbol{v}_o\cdot\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \boldsymbol{v}_o\cdot\frac{1}{c^2}\frac{\partial(\nabla\phi)}{\partial t} = \left(\frac{mc}{\hbar}\right)^2(f\phi + \boldsymbol{v}_o\cdot\boldsymbol{A}). \tag{7.7}$$

As f is an arbitrary function, it can be defined in order that

$$f\frac{\partial(\nabla\cdot\boldsymbol{A})}{\partial t} - \boldsymbol{v}_o\cdot\frac{1}{c^2}\frac{\partial(\nabla\phi)}{\partial t} - \boldsymbol{v}_o\cdot\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} = -f\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}; \tag{7.8}$$

hence

$$f\nabla^2\phi - f\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \left(\frac{mc}{\hbar}\right)^2(f\phi + \boldsymbol{v}_o\cdot\boldsymbol{A}). \tag{7.9}$$

Eventually it is possible to infer from this equation

$$f\nabla^2\phi - f\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \left(\frac{mc}{\hbar}\right)^2(1+q)f\phi, \quad \boldsymbol{v}_o\cdot\boldsymbol{A} = qf\phi, \quad q \gtreqless 0; \tag{7.10}$$

both f and $\boldsymbol{v}_o$ are arbitrary, thus the last position just introduced is in fact possible. Actually the first (7.10) does not depend on f and has still the form of a Klein Gordon-like equation for the previous scalar function ϕ, where appears however the factor $1+q$ at the right hand side corresponding to the previous $\sigma(1\pm\sigma)$ of (7.3). Just this factor is the key to split this double valued equation: indeed $q=0$ is one scalar equation, whereas the additional chance $q\neq 0$ allows inferring the couple of Proca equations via the position $\boldsymbol{v}_o\cdot\boldsymbol{A} = qf\phi$ that introduces the vector field $\boldsymbol{A}$. It is enough to revert the steps from (7.10) to (7.5) still via the key position (7.8). The profound physical meaning of this position, here purposely introduced to obtain (7.9), is shortly outlined below, to show that it is not merely a useful algebraic step.

The Equation (7.8) reads

$$f\left(\frac{\partial(\nabla\cdot\boldsymbol{A})}{\partial t} + \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}\right) = \frac{\boldsymbol{v}_o}{c^2}\cdot\left(\frac{\partial(\nabla\phi)}{\partial t} + \frac{\partial^2\boldsymbol{A}}{\partial t^2}\right) \tag{7.11}$$

i.e.

$$f\frac{\partial}{\partial t}\left(\nabla\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t}\right) = \frac{\boldsymbol{v}_o}{c^2}\cdot\frac{\partial}{\partial t}\left(\nabla\phi + \frac{\partial\boldsymbol{A}}{\partial t}\right). \tag{7.12}$$

In this equation it is possible to put formally

$$\nabla\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t} = 0, \quad \nabla\phi + \frac{\partial\boldsymbol{A}}{\partial t} = -\boldsymbol{E}, \quad \frac{v_o}{c^2}\cdot\frac{\partial\boldsymbol{E}}{\partial t} = 0 \tag{7.13}$$

the first position is the Lorentz condition, the second equation is the definition of the new quantity $\boldsymbol{E}$, the third equation is obtained replacing $\boldsymbol{E}$ at the right hand side of (7.11) and simply means that $\boldsymbol{v}_o$ is orthogonal to $\partial\boldsymbol{E}/\partial t$ to get $0=0$ at both sides. Now give $\boldsymbol{A}$ physical meaning introducing the following positions

$$\boldsymbol{H} = \nabla\times\boldsymbol{A}, \quad \nabla\cdot\boldsymbol{H} = 0, \quad \nabla\times\boldsymbol{E} = -\frac{1}{c}\frac{\partial\boldsymbol{H}}{\partial t} \tag{7.14}$$

the first position is simply the definition of a new quantity $\boldsymbol{H}$, whence follows the second equation by consequence; the third equation is obtained taking $\nabla\times$ of both sides of the second (7.13). Note now that the first and second (7.13) yield

$$-\nabla\cdot\left(\nabla\phi+\frac{\partial\boldsymbol{A}}{\partial t}\right)=-\left(\nabla^2\phi+\frac{\partial\nabla\cdot\boldsymbol{A}}{\partial t}\right)=\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}-\nabla^2\phi=\nabla\cdot\boldsymbol{E}=\rho \qquad (7.15)$$

Eventually it is possible to infer a fourth equation considering the continuity condition $\delta g=0$ of an arbitrary function $g=g(x,t)$, which reads $\delta g=(\partial g/\partial t)\delta t+(\partial g/\partial x)\delta x=0$; thus the one dimensional expression $\partial g/\partial t+(\partial g/\partial x)v=0$, where v must be intended of course as $v_x=\delta x/\delta t$, reads in general $\boldsymbol{v}\cdot\nabla g=-\partial g/\partial t$. Hence $\boldsymbol{v}\cdot\nabla g=\nabla\cdot(g\boldsymbol{v})-g\nabla\cdot\boldsymbol{v}$ yields

$$g\nabla\cdot\boldsymbol{v}=0=\frac{\partial g}{\partial t}+\nabla\cdot(g\boldsymbol{v}).$$

Since g has not yet been defined, it is possible to rewrite this equation implementing the scalars already inferred, to obtain a self-contained set of equations. Put

$$g=\frac{1}{c}\nabla\cdot\boldsymbol{E}=\rho,\quad g\boldsymbol{v}=\rho\boldsymbol{v}=\boldsymbol{J},$$

which yields therefore

$$\nabla\cdot\nabla\times\boldsymbol{X}=0=\frac{1}{c}\frac{\partial\nabla\cdot\boldsymbol{E}}{\partial t}+\nabla\cdot\boldsymbol{J};$$

It is convenient to utilize the vector property $\nabla\cdot\nabla\times(\,)=0$ in agreement with $\nabla\cdot\boldsymbol{v}=0$ of solenoidal flux of ρ to simplify this last equation whatever the arbitrary field $\boldsymbol{X}$ might be. Thus one obtains

$$\nabla\times\boldsymbol{X}=\frac{1}{c}\frac{\partial\boldsymbol{E}}{\partial t}+\boldsymbol{J},\quad \boldsymbol{J}=\rho\boldsymbol{v}. \qquad (7.16)$$

Despite the lack of specific information to identify the actual nature of $\boldsymbol{X}$, it is reasonable to put $\boldsymbol{X}\propto\boldsymbol{H}$ with $\boldsymbol{H}$ magnetic field to be introduced just in the next section: this position is in fact possible without introducing a new field, hardly justifiable in the present context. Hence the position (7.8) is valuable as it implies four relevant equations, whose importance appears by answering questions like: flux of what? what kind of fields are $\boldsymbol{E}$ and $\boldsymbol{H}$? The next section clarifies these points that clearly allow to obtain the Maxwell equations (7.14), (7.15) and (7.16).

These fields allow modifying appropriately the functional (7.1), in order to describe one particle even via a possible interaction potential; of course such a calculation is omitted here for brevity, being clearly beyond the purposes of the present paper.

8. The Fundamental Interactions of Nature: Force Laws

Some considerations about the gravity force in the subsection 6.2 were inferred starting from the Equation (4.9). Now the concept of force is reexamined starting from the more general Equations (3.1) and (3.69). This section consists of three subsections.

8.1. Preliminary Considerations

The Equations (3.1) provide several chances of defining the general concept of force, directly related to (3.69) and (3.70)

$$\frac{\delta p}{\delta t}=\frac{\delta\varepsilon}{\delta x}=F \tag{8.1}$$

whence, implementing once more $\delta p=\hbar/\delta x$, one infers in general

$$\delta\dot{p}=-\frac{\hbar}{\delta x^2}\delta\dot{x},\quad \delta\dot{x}=\frac{\delta x}{\delta t},\quad \delta\dot{p}=\frac{\delta p}{\delta t}. \tag{8.2}$$

To understand intuitively the correlation between F and $\delta\dot{x}$, think that the effect of any force is to modify the state of motion of matter on which it acts. Consider a body of mass m under the action of a force F. If m is actually a system of particles, then F perturbs the dynamics of all particles of m: for example they move faster. According to (3.1), modifying the kinetic energy of a system of particles implies changing in principle their delocalization extent and thus the range size δx able to include each one of them. The greater the force altering the status of a system, the greater must be the size change rate $\delta\dot{x}$ to account for the altered delocalization extent of matter in δx. This is the intuitive way to justify in general the quantum link between F and the related $\delta\dot{x}$. More specifically (8.1) also imply

$$v=\frac{\delta x\delta\varepsilon}{\hbar},\quad \delta x\delta\varepsilon\le\hbar c; \tag{8.3}$$

the inequality is direct consequence of the first equation with δt^{-1} replaced by $\delta\varepsilon/\hbar$. Again, the position $\delta\to\partial$ yields the usual definition of generalized local force $F=\dot{p}$. These equations are directly referable to long range interactions, because δx is defined even at infinity in lack of specific constrains; being directly inferred from the general Equations (3.1), F is expected to hold for charged and neutral particles. In effect the Newton and Coulomb forces represent an important class of forces that vanish at infinity as δx^{-2}, justifiable simply assuming $\delta\dot{x}=const$; is evident the analogy of (8.2) with F_G of (6.10).

Beside (8.2), particularly interesting are further dimensional considerations about characteristic space and time ranges related some specific forces. Write for example

$$F=\frac{\hbar^2}{mV}=\frac{(\hbar c)^2}{mc^2V}=\frac{\hbar}{V}\frac{\delta x^2}{\delta t},\quad F=\frac{\hbar}{\delta x\delta t} \tag{8.4}$$

the first definition follows from $\delta p^2/m\delta x=\hbar^2/m\delta x^3$. The Equations (8.4) and (8.3) prospect the chance of introducing the concept of short range force as that related to characteristic lengths, times and possibly volume consistently with the Equations (3.1). An example in this respect is the zero point energy resulting from the confinement of matter in a finite volume of space time, already introduced in (5.23); another example is the Casimir force, shortly sketched below.

The ranges $\delta\varepsilon$ and δx of (8.4) help to fix the energy scale or the distance scale characteristic of specific interactions. Are useful in this respect the values (3.12) of length ranges; the Equations (8.4) are implementable with these scale lengths of prospective interest to estimate the strength of short range forces.

The few remarks exposed here highlight how to proceed in various cases. First of all, it is necessary to examine how long range and short range forces are included in the general definition (8.1) of *F*: the key point is the deformation rate $\delta\dot{x}$ of the space time range δx.

Let two interacting partners be δx apart and expand in series $\delta\dot{x} = \delta\dot{x}\left(\delta x^{-1}\right)$: this position ensures that *F* vanishes at $\delta x \to \infty$. Write thus in general

$$\delta\dot{x} = \delta\dot{x}_0 + \frac{Ac}{\delta x} + \frac{Bc}{\delta x^2} + \cdots \tag{8.5}$$

where *A* and *B* are the constant coefficients of the series expansions. So (8.2) splits into the sum of various terms

$$F_0 = \pm\frac{\hbar\left|\delta\dot{x}_0\right|}{\delta x^2}, \quad F_1 = \pm\frac{\hbar\left(\ell_0 + \ell_1\right)c}{\delta x^3}, \quad F_2 = \pm\frac{\hbar c\left(\ell_2 + \ell_3\right)^2}{\delta x^4},$$
$$\ell_0 + \ell_1 = A, \quad \left(\ell_2 + \ell_3\right)^2 = B,$$

Here $\delta\dot{x}_0 = const$ by definition, with signs of F_0 dependent on whether $\delta\dot{x}$ swells or shrinks during δt. The Equation (8.5) is more general than (6.7): the latter concerns specifically $\mathcal{G}$ and thus the gravity, *i.e.* F_0 corresponds to F_G, whereas (8.5) instead concerns more in general $\delta\dot{x}$; the higher order terms expressing $\delta\dot{x}$ include $\mathcal{R}$ of (6.7). Just to show this point the coefficients *A* and *B* have been split introducing constants lengths ℓ_i that characterize various kinds of forces in fact included in (8.2) and made explicit by the respective terms of the series expansion of *F*. Examine thus one by one the forces resulting from the first three terms of (8.5) to show that in effect this approach is interpretable in a sensible way. The first force identifiable is

$$F_{NC} = \pm\frac{\hbar\left|\delta\dot{x}_0\right|}{\delta x^2} + \frac{\hbar c\ell_1}{\delta x^3} + \frac{\hbar c\ell_2^2}{\delta x^4} + \cdots, \tag{8.6}$$

where the subscript *NC* stands for Newton Coulomb. The second and third forces easily identifiable are

$$F_{zp} = \frac{\hbar\ell_0 c}{\delta x^3} = \pm\frac{\hbar^2}{2mV_{zp}}, \quad \ell_0 = \frac{\hbar}{2mc},$$

having inferred ℓ_0 by comparison with (5.23), and

$$P_{Ca} = \frac{F_2}{\ell_2\ell_3} = \pm\frac{\hbar c\sigma}{\delta x^4}, \quad \sigma = \frac{\ell_3}{\ell_2} + 2.$$

One finds again the zero point force controlled by the Compton length ℓ_0 of the mass *m*. Moreover one recognizes the Casimir force per unit surface $\ell_2\ell_3$ given by $\hbar c$ times the pertinent numerical factor σ whose value is controlled by ℓ_3 whatever the value of ℓ_2 contributing to F_G might be. In this way neither ℓ_0 nor ℓ_1 and ℓ_2 result fixed; so the second and third terms of (8.6) are the higher order terms of the zero order approximation F_0 in principle definable by these lengths independently of the coefficients characterizing F_{zp} and force per unit surface P_{Ca}. The zero point energy has been already concerned; a detailed discussion of the Casimir term is clearly outside the purposes of the present paper. It is really crucial the fact that various kinds of forces are nested in the general uncertainty Equation (8.2), in turn direct consequence of (2.10). Owing to the importance of (8.6), the following consideration will be focused on this equation for sake of brevity only; this allows to complete the information in Section 6.2.

8.2. Long Range Gravity Force

To verify how (8.6) implies more specifically the space time curvature, examine both sides of the general Equation (8.2). The Equations (3.1) imply the following chain of equations implementing the left hand side of (8.2)

$$\frac{\delta p}{\delta t}=\frac{\hbar}{\delta t}\frac{1}{\delta x}=\frac{\hbar}{\delta t}\frac{\delta x'}{\delta x'\delta x}=\frac{\hbar}{\delta t}\frac{\delta x'}{\delta x+\delta x'}\left(\frac{1}{\delta x}+\frac{1}{\delta x'}\right)=\frac{\delta p'}{\delta t}r_c\delta x' \qquad (8.7)$$

$$\delta p'=\frac{\hbar}{\delta x+\delta x'},\quad r_c=\frac{1}{\delta x}+\frac{1}{\delta x'},$$

being r_c the Laplace average curvature radius of an ideal surface such that $F\to 0$ for $\delta x\to\infty$ and $\delta x'\to\infty$.

Note that r_c is formally similar to $K_L=1/r_1+1/r_2$, where the addends are defined on two orthogonal planes called curvature sections; it refers to flat space time. According to its classical derivation, the local value of K_L is calculable as both radii r_1 and r_2 are assumed in principle exactly knowable. Here, instead, the quantum derivation of r_c does not allow any information about size and even mutual orientation of δx and $\delta x'$, the only indication available being that it is conceptually defined by two curvature sections in a 4D space time; so r_c is not calculable in practice, it is compatible with all combination of values included within δx and $\delta x'$. The relativistic curvature K_E is instead self-defined in a Gauss curvilinear coordinates in a covariant way regardless of the reference system. According to (3.2), however, even r_c defined by two uncertainty ranges actually waives the link to a specific reference system; hence the impossibility of calculating uniquely r_c prevents comparing it to K_E, whose local value is instead in principle calculable. Hence is meaningless to enquiry whether or not the Equation (8.7) fits the standard definition of general relativity or not. In other words K_E must be necessarily covariant to have physical meaning, because effectively it can be calculable locally; r_c instead fulfills the quantum concept of covariance required by (3.2) and cannot take any deterministic value. Hence, in lack of numerical assessment, the quantum reasoning implied by (8.7) allows conceptual comparison only: the quantum definition of space time curvature, although symbolic, is still related to its relationship with F, *i.e.* $F \neq 0$ for δx and $\delta x'$ both finite. The non trivial implication of this reasoning is that in this way the relativity becomes a corollary of (3.1) and (8.2) via the series expansion (8.5) that generalizes the mere Newtonian term. This is not surprising because the Section 1 has evidenced the 4D holistic character of the present model; the purpose of the next considerations is to clarify further this idea.

Recalling (2.36) $p = h/\lambda$, the right hand side of (8.2) yields, owing to (4.24) and (2.36),

$$-\frac{\hbar}{\delta x}\frac{\delta \dot{x}}{\delta x} = -h\frac{\delta\dot{\lambda}}{\delta\lambda^2} = \mp\frac{\delta p}{\delta t}; \tag{8.8}$$

so, merging (8.7) and (8.8), one finds

$$-\frac{h\delta\dot{\lambda}}{\delta\lambda^2} = \mp\frac{\delta p'}{\delta t}r_c\delta x' = \mp r_c\delta\varepsilon', \quad \delta\varepsilon' = v'\delta p', \quad v' = \frac{\delta x'}{\delta t}. \tag{8.9}$$

No apparent reference to the mass is explicitly evident in this formula: the force $r_c\delta\varepsilon'$ and its approximation F_0 are due to the mere deformation rate of the range $\delta\dot{x}_0$.

On the one hand (8.2) results consistent with this equation that links force $\delta\dot{p}$ and energy $\delta\varepsilon'$ via the curvature radius r_c of space time: as expected the momentum, and thus its time change as well, can be expressed via corpuscle formalism, see e.g. (3.64), and via wave formalism inferred in (2.36). Accordingly $F = \dot{p}$ waives the concept of mass if just (2.36) is implemented to calculate $\dot{p} = -h\delta\dot{\lambda}/\delta\lambda^2$. The analogy with (8.6) and (8.5) appears because also now it is possible to write

$$\delta\dot{\lambda} = \pm\delta\dot{\lambda}_0 + \sum_j k_j\delta\lambda^{-j} \tag{8.10}$$

this is the meaning of the Equations (8.7) to (8.9), where the mass is mere dimensional property of *h*.

On the other hand, the connection of (8.6) with (6.12) implied by (6.7) requires writing

$$F_0 = \pm G\frac{m_1m_2}{\delta x^2}, \quad \delta\dot{x}_0 = m_1m_2\frac{G}{\hbar} \tag{8.11}$$

the curvature r_c is linked to one mass, that defining ε' and $\delta\dot{p}$ of (8.9), the zero order deformation rate $\delta\dot{x}_0$ of δx is given by the constant $G/\hbar$ times the product m_1m_2, *i.e.* $\delta\dot{x}_0$ is proportional to m_1m_2 that in turn is directly proportional to the force. The Equations (8.6) and (8.11) are the space time/matter formulations of the gravity force analogous to the wave/corpuscle formulations of the energy and momentum in quantum mechanics; in effect (3.70) has shown that in general the force is proportional to $\delta\varepsilon\delta p$.

It turns out therefore from the previous considerations with the help of (2.36)

$$\left|\frac{\delta p}{\delta t}\right|_{Ein} = |F| = \left|\frac{\delta p}{\delta t}\right|_{New} \leftrightarrow r_c\delta\varepsilon' = \frac{h\delta\dot{\lambda}}{\delta\lambda^2} = G\frac{m_1m_2}{\delta x^2}, \tag{8.12}$$

whence the correlations

$$\text{force field} \rightarrow \text{deformation rate of space time ranges} \rightarrow \text{acceleration} \qquad (8.13)$$

Comparing the left and right hand sides of both (8.12), is evident why Einstein has successfully replaced the concept of force with that of space time curvature, while skipping the more intuitive Newtonian correspondence between mass and force: thus, in His intuition, the mass appears directly related to the space time curvature. So the first correlation (8.13) is understandable. The second correlation deserves attention, as the concept of acceleration has been not yet introduced; it will be concerned in the next subsection in particular to explain what have to do $\delta\dot{x}$ or $\dot{\lambda}$ with the acceleration implied by *F*. It will be shown that just the ranges, which link the concept of force to the quantum uncertainty, also plug the Newton and Coulomb forces into the realm of quantum mechanics.

8.3. The Equivalence Principle

Implement the Equations (8.11) and (8.8) to understand why the mass appears in the former and not in the latter. Also this topic, shortly sketched here for completeness, has been concerned in [10].

Position and size of any $\delta x = x - x_o$ in an arbitrary R are respectively definable considering the distance of either range boundary, say x_o, from the origin O and the distance of x from x_o; of course the opposite choice would be identically admissible. Being both boundaries arbitrary, in general it is possible to regard $x = x(t)$ and $x_o = x_o(t)$. A force F arises inside δx because in general $\dot{x} \neq \dot{x}_o$, *i.e.* when the range size of δx shrinks or stretches as a function of time during $\delta t = t - t_o$ with respect to its initial size at the time t_o. To simplify the reasoning it is enough to examine the cases where: 1) $x_o = const$ and $x = x(t)$ only or 2) $x = const$ and $x_o = x_o(t)$ only; as anyway the size of δx changes, for example because of energy fluctuation of a particle inside δx, both cases imply in general $F \neq 0$ and the following considerations about inertial and accelerated reference systems.

Imagine an observer sitting on x_o and assume for simplicity that one particle only is delocalized in δx; the rising of any F is detected observing the dynamical behavior of this test particle.

In the case (1) the observer is by definition at rest in R with respect to *O*; yet he acknowledges a force $F = -\left(\hbar/\delta x^2\right)\delta\dot{x}$ acting on the particle, actually due to $\dot{x} \neq 0$. To justify in principle why the motion of the particle is perturbed, the observer reasonably thinks to the presence of an external force, e.g. a gravitational mass outside δx.

In the case (2) the observer no longer at rest in R necessarily accelerates with respect to *O*, whereas the force $F_o = \left(\hbar/\delta x_o^2\right)\delta\dot{x}_o$ again appears in δx according to $\dot{x}_o \neq 0$; now F_o governs the dynamics of the particle delocalized in δx. The observer concludes that its own acceleration is due to F_o.

Of course the analytical forms of F and F_o are in principle analogous, although their strengths are in general different if $\dot{x} \neq \dot{x}_o$; indeed the forces only differ by the time dependence of either boundary coordinate of δx with which is calculated the overall $\delta \dot{x}$. However, despite the boundaries of δx are in general arbitrary and independent each other, nothing hinders to assume in particular $\dot{x}_o / x_o^2 = -\dot{x}/x^2$: *i.e.* stretching of δx occurs via forwards displacement of x only or backwards displacement of x_o only with respect to *O*. So locally $F = -F_o$. The acceleration experienced by the observer and the presumed force F arising outside δx perturb in the same way the test mass because they have actually a unique background, the deformation rate (8.5) of the space time range δx itself that in effect implies (8.6): in (2) this deformation rate is perceptible by the observer as force F_o whereas in (1) it does not, although in both cases the observer can anyway record the same change of local dynamics of the particle inside δx. Clearly the observer reference system R_o with origin on x_o is at rest or inertial in (1) but non-inertial in (2) with respect to *R*, in agreement with the aforesaid correlation.

One key point of the reasoning is that these conclusions hold exchanging the role of x and x_o, because both range boundaries are arbitrary and physically equivalent; no physical property characterizes specifically either space time boundary displacement. As (1) and (2) are physically indistinguishable, the unique information available is the overall deformation rate $\delta \dot{x}$ and its related *F*; this means that the concept of acceleration implied by F_o holds identically for F as well. Another key point is that the acceleration does not necessarily imply the concept of mass, but that of force in turn due to inertial and non-inertial reference systems.

One could also say that the concept of force is redundant, what in fact exists is the stretching/shrinking rate of δx which in turn implies space time curvature according to (8.7). But now this statement has quantum foundation only.

Anyway the dualism wave/corpuscle of quantum mechanics has relativistic analogy in the “corpuscular” Equation (6.9) and “wave” Equation (8.9) properties of matter; the latter originates from the space time curvature, the former from the necessity of defining the change of delocalization extent of massless or massive particles both contextually implied by the probabilistic Equation (4.7). In this sense quantum and relativistic physics are perfectly symmetrical, which is not surprising because both are rooted on the quantum uncertainty. From this analogy follow the correspondences (8.12) along with the Equations (6.28) and the various forces implied by (8.5).

Is evident the analogy of δx with the elevator of Einstein’s thought experiment: the cases of inertial and non-inertial reference systems merge here into the unknown and unknowable behavior of the boundaries of a unique space time uncertainty range δx only. Implementing space time ranges rather than local space time coordinates plugs a typical relativistic reasoning about inertial and non-inertial reference systems into the quantum uncertainty (3.1). In this sense the relativity is conceptually compatible with quantum requirements; any reasoning via local coordinates, e.g. the tensor calculus, wouldn't. The present approach shows what the mere wave formalism to quantum mechanics cannot emphasize itself: quantum and relativistic theories are conceptually rooted in the unique concept of uncertainty, the operator formalism exposed in the subsection 3.5 is instead less general being actually itself a corollary of the quantum uncertainty.

8.4. Long Range Electromagnetic Interactions

Start from the Equation (4.8) $\Delta\varepsilon = \mathrm{n}^2\delta\varepsilon$ and the position (4.9) $\Delta\varepsilon = \hbar/\Delta t$, which now are once more under test besides to the results (4.13), (4.14), (4.16) and (4.17) already obtained; the Equation (4.9) was also the starting point of the section 6.2. Recalling the definition (4.5) of refractive index n, elementary manipulations yield

$$\delta\varepsilon = \frac{v^2}{c}\frac{\hbar}{\ell} = \varepsilon_2 - \varepsilon_1, \quad \ell = c\Delta t, \tag{8.14}$$

being of course ℓ an arbitrary length. The second equality is the mere definition of energy range $\delta\varepsilon$ with arbitrary boundary values ε_2 and ε_1. A possible way to split accordingly $v^2\hbar/\ell c$ too, is to rewrite (8.14) defining ε_1 and ε_2 as follows

$$\delta\varepsilon = \frac{\varepsilon_1^2}{\varepsilon_2}\frac{q^2}{\hbar c}, \quad \varepsilon_1 = \frac{\hbar v}{\ell}, \quad \varepsilon_2 = \frac{q^2}{\ell} \tag{8.15}$$

to obtain next (8.14) rewritten as

$$\frac{v^2}{c}\frac{\hbar}{\ell} = \varepsilon_2 - \varepsilon_1 = \frac{q^2 - \hbar v}{\ell} \tag{8.16}$$

the positions (8.15) convert thus the unique term at the left hand side of (8.14) into the difference of energies q^2/ℓ and $\hbar v/\ell$ defining $\delta\varepsilon$. In general $q = q(v)$. Let ℓ measure the distance between two interacting partners. The fact that ℓ is defined by c means that the carriers of the force are massless particles, photons. Also, v that defines the energy range (8.14) characterizing this kind of interaction must be consistently identified by $v = c$; as any $v < c$ could not be enough to travel through ℓ, a coherent way to characterize the peculiar value of q consistent with ℓ requires the boundary condition $\mathrm{n} = 1$ in (8.16). Put thus the resulting value of q, now uniquely defined, proportional to a "new" quantity called e. So, calling $2\alpha_0$ the dimensionless proportionality constant linking q^2 and e^2, (8.16) yields

$$\frac{\hbar c}{\ell} = \frac{2\alpha_0 e^2 - \hbar c}{\ell}, \quad q^2 = 2\alpha_0 e^2, \quad v = c. \tag{8.17}$$

Multiplying both sides by $\ell/\hbar c$ one finds

$$\alpha_0 = \frac{1}{\alpha}, \quad \alpha = \frac{e^2}{\hbar c}$$

as a result we have obtained the definition of fine structure constant α via three universal constants of the nature, whereas α_0 is to be regarded as a pure number corresponding to the numerical value of $1/\alpha$.

Note in this respect that the electric charge *e*, so far not yet explicitly introduced, appears in the model via α. This introductory reasoning outlines the next task to be concerned just now: to show how the Equation (4.9) implies the electromagnetic interaction too. This subsection links therefore the following considerations to the Section 7. Usually *e* is introduced by postulating the Coulomb force; here instead *e* and thus the electromagnetic forces are introduced starting from α. By analogy with (8.6), the Equation (8.2) is rewritten as follows

$$\pm\frac{\hbar c}{\delta x^2}=F_C=\pm\alpha^{-1}\frac{e^2}{\delta x^2},\quad \delta\dot{x}_0=c, \tag{8.18}$$

i.e. the physical meaning of α introduces itself also the long range Coulomb force component F_C.

Is evident the formal analogy between F_0 of (8.11), concerning the mass linear density $m/\delta x$, and F_C of (8.18), introducing the charge linear density $e/\delta x$: both regard in particular $|\delta\dot{x}_0|\times\hbar$ as constant characterizing the lowest order term (8.6) of the series expansion (6.7). This suggests that the first order approximations of Newton and Coulomb laws should be both deducible from $|\delta\dot{x}_0|\times\hbar$ through a dimensional constant. In effect the connection between G and *e* via a dimensional proportionality constant is easily proven; indeed

$$|e|=(\xi_{eG}\alpha)G$$

yields numerically

$$e=4.80\times10^{-10}\ \text{ues},\quad \alpha G=4.88\times10^{-10}\ \text{cm}^3/\text{s}^{-2}\cdot\text{g},$$
$$\xi_{eG}\approx1.01\ \text{ues}\times t^2\ \text{g}/\text{cm}^3=(\text{g}/\text{cm})^{3/2}\cdot\text{s}.$$

Of course the signs of the component F_C correspond to equal or opposite charges defining α. Implement first (8.18) for both charges at rest in *R*; so

$$F_C=\frac{e_r^2}{\delta x^2},\quad v=0, \tag{8.19}$$

where the subscript stands for "rest"; in this equation there is no explicit reference to *v*, which from now on denotes the relative motion of the charges. As δx is arbitrary, α has been included in it, to simplify notations like $\delta x^*=\sqrt{\alpha}\delta x$.

The second way to implement (8.18) assumes constant the rate *v* with which moves either charge with respect to the other at rest in *R*; as according to (2.34) $\delta p=\hbar/\delta x=(v\delta\varepsilon+\varepsilon\delta v)/c^2$, (8.18) yields

$$\frac{e^2}{\delta x^2}=\frac{\hbar c}{\delta x^2}=\frac{\delta(pc)}{\delta x}=\frac{1}{c}\left(\frac{v\delta\varepsilon}{\delta x}+\frac{\varepsilon\delta v}{\delta x}\right). \tag{8.20}$$

Thus

$$\frac{e_m^2}{\delta x^2}=\frac{v}{c}\frac{\delta\varepsilon}{\delta x},\quad v=const,\tag{8.21}$$

where the subscript stands for "mobile", whence

$$\frac{e_m^2}{\delta x^2}=\frac{\delta\varepsilon}{c\delta t}=\frac{\delta\varepsilon\sqrt{1-v^2/c^2}}{c\delta t'}.\tag{8.22}$$

The last equality has introduced the time range $\delta t_r'$ defined in a reference system R' of the mobile charge moving at rate $-v$ with respect to *R*; an observer ideally sitting on the moving charge in R' sees the other charge at rest, as the backwards motion of R' in *R* balances exactly the forwards motion of the charge. So, in practice (8.22) can be regarded likewise (8.19); obviously (8.22) reduces anyway to (8.19) in particular for $v=0$. Hence it is possible to write the second equality as

$$F_{em}=\frac{\delta\varepsilon}{c\delta t'}=\frac{e^2}{\delta x'^2\sqrt{1-v^2/c^2}},\quad F_{Lor}=vH\frac{e}{c};\tag{8.23}$$

the first equation emphasizes the link between $e^2/\delta x^2$ in R and R', the second equation also follows directly from (8.21) according to the following chain of elementary steps

$$\frac{v}{c}F=\frac{v}{c}e\frac{F}{e}=\frac{e}{c}vH,\quad F=\frac{\delta\varepsilon}{\delta x},\quad H=\frac{F}{e}.$$

Regarding $F=F_{em}$ one finds a "new" quantity called magnetic field already introduced as a final step (7.16) of the reasoning in the section 7. The one dimensional scalar approach followed throughout this paper hides the actual vector character of $\boldsymbol{v}$ and thus of $\boldsymbol{F}_{em}$ and $\boldsymbol{F}_{Lor}$. Simple considerations allow however to surrogate this missing information acknowledging that F_{Lor} and F_{em} are two different corollaries of a unique information to describe the charge dynamics: both Equations (8.23) follow from different ways of rewriting the unique Equation (8.21), as in effect it is physically sensible. So, to avoid that the energy of a mobile test charge in both fields is counted twice summing separately F_{em} and F_{Lor}, it is necessary that the former only performs work on the mobile charge, likewise as in the particular case of charge at rest, whereas the latter doesn't; the vector properties of these forces follow from these considerations, *i.e.* $\boldsymbol{H}$ and $\boldsymbol{v}$ must be such that $\boldsymbol{F}_{Lor}\propto\boldsymbol{H}\times\boldsymbol{v}$.

The third way to handle α concerns the case where v is not constant; owing to (8.20), an additional force term F_ω is expected because of the addend $\varepsilon\delta v/c\delta x$ previously omitted in (8.21). This additional term reads

$$F_\omega=\frac{\varepsilon\delta v}{c\delta x}\frac{\delta t}{\delta t}=\frac{\varepsilon\delta t}{\delta x}a=\frac{\varepsilon\hbar}{\delta x\delta\varepsilon}a,\quad a=\frac{1}{c}\frac{\delta v}{\delta t},$$

where the acceleration a of the charge appears as a reciprocal frequency because of the factor c^{-1}. It is possible to extract from this chain of equation the frequency ω defined by the energy $F_\omega\delta x$, *i.e.*

$$\frac{F_\omega\delta x}{\hbar}\frac{\delta\varepsilon}{\varepsilon}=\omega=a\tag{8.24}$$

A fraction $\delta\varepsilon/\varepsilon$ of energy $F_\omega \delta x$ is thus converted into and appears as electromagnetic radiation, whose energy $\hbar\omega$ increases of course with $\delta\varepsilon/\varepsilon$. The fact of having found ω for $\delta v \neq 0$ means that an accelerated charge implies emission of e.m. radiation. Of course $F_\omega = 0$ and thus $\omega = 0$ for $v = 0$ or $v = const$, so even this contribution to the right hand side of (8.20) vanishes for charges at rest. In principle $\varepsilon\omega$ prospects the chance of calculating the power irradiated by an accelerated charge regarding appropriately F_ω and $\delta\varepsilon$ of (8.24). This chance has been exemplified in section 7 in a more complete and rigorous way via the Maxwell equations.

Is still proven useful here the initial idea of implementing uniquely the early Equations (1.11) and (1.12) without any further physical hint but simply including α among the fundamental constants of Nature.

8.5. Quantum Charges

Dividing both sides of the inequality (8.3) by e^2 and next by δx too, one finds

$$\delta\varepsilon = \frac{\sigma^2}{\alpha}\frac{e^2}{\delta x}, \quad \sigma^2 \leq 1, \tag{8.25}$$

being σ an appropriate factor. According to (3.69) and (3.70) this yields also

$$F_{e'} = \frac{\delta\varepsilon}{\delta x} = \frac{1}{\alpha}\frac{e'^2}{\delta x^2}, \quad e' = \sigma e, \tag{8.26}$$

Now the question rises: does e' have mere numerical meaning or it actually generalize the concept of usual charge e according to $e' \leq e$? To answer this question consider first some implications of (8.26) based on (3.70). As $F_{e'} = PA$, where A is an arbitrary area and P pressure, one finds the dimensional relationships

$$P = \frac{F_{e'}}{A} = \frac{1}{\alpha}\frac{e'^2}{A\delta x^2} \tag{8.27}$$

the Equations (8.26) imply the rising of a pressure P related to the energy density ε/V due to charged particles enclosed in the volume V. Start thus just from the dimensional identity between pressure and energy density and write

$$P = \sigma'\frac{E}{V}, \quad V = \Delta x^3 = \sigma' A\delta x, \tag{8.28}$$

where σ' is a proportionality factor necessary to define in general the volume V as a function of A defining the pressure. Owing to the dimensional character of this equation, although P and E have been defined specifically by (8.27) and (8.28), the following reasoning holds in general for any E, *i.e.* also for atoms, ions, elementary charges and even photons. The second equation yields

$$\frac{\delta V}{V} = 3\frac{\delta\Delta x}{\Delta x}. \tag{8.29}$$

Consider now that if in V are contained photons or matter, e.g. gas particles, it is possible to implement the wave properties of matter and write

$$V=\left(\frac{v}{\nu_{\lambda}}\right)^{3}=\left(\frac{c}{\nu}\right)^{3},\quad \nu=\nu_{\lambda}\frac{c}{v}=\frac{c}{\lambda},\quad \lambda=\frac{v}{\nu_{\lambda}},\quad \nu\geq\nu_{\lambda} \tag{8.30}$$

the matter particles are assumed moving at average rate v with De Broglie momentum h/λ and thus frequency ν_{λ}; for calculation purposes, ν_{λ} has been rewritten as a function of ν as indicated here to include also photons. The presence of steady waves in V requires

$$\frac{\delta V}{V}=-3\frac{\delta\nu}{\nu}=-3\frac{\delta E}{E} \tag{8.31}$$

the first equality is directly deducible from (8.30) that expresses the steadiness condition, *i.e.* the change of λ requires that of Δx as $\lambda=n\Delta x$ with n integer, the second equality expresses the proportionality between ν and E. So the change of E inside V is related to that of ν of the matter/light waves propagating in V. Implement now the idea that (8.29) regards a number of corpuscles inside V, whose change of energy density is uniquely definable by δV anyhow it might be obtained; instead (8.31) regards waves, whose energy changes are presumably related to how the early V is modified by a given δV because of steadiness condition. Reasonably the steady propagation of waves is different depending on whether one side only or two sides or even three sides of V are modified. Consider thus the three possible ways to deform the initial V, whose ΔV remains however uniquely defined in all cases: V can be equivalently rewritten as $\Delta x_0^2\Delta x$ or $\Delta x_0\Delta x^2$ or Δx^3. Being both Δx and the constant Δx_0 arbitrary, it is certainly possible to define them in order to fit a given value V of course arbitrary itself. Hence

$$\frac{\delta V}{V}=n_{\nu}\frac{\delta\Delta x}{\Delta x},\quad n_{\nu}=1,2,3.$$

To make consistent both ways of defining $\delta V/V$ merge this result with (8.31) to obtain

$$n_{\nu}\frac{\delta\Delta x}{\Delta x}-3\frac{\delta E}{E},$$

whence

$$\frac{\delta E}{\delta\Delta x}=-\frac{n_{\nu}}{3}\frac{E}{\Delta x},\quad \frac{n_{\nu}}{3}=\frac{1}{3},\frac{2}{3},1. \tag{8.32}$$

Assume that the left hand side defines an average force such that $\delta E=\langle F\rangle\delta\Delta x$; dividing both sides by the surface Δx^2 one finds

$$\frac{\langle F\rangle}{\Delta x^{2}}=P=-\frac{n_{\nu}}{3}\frac{E}{\Delta x^{3}}. \tag{8.33}$$

At the right hand side appears an energy density ρ_E defined by an amount of energy δE, arbitrary, in the volume $\Delta x^2\delta\Delta x$, arbitrary as well, and recall the initial position (8.28); then, as the dimensions of $E/\Delta x$ and Δx^2 are force and surface, one finds eventually

$$P_1=\frac{1}{3}\rho_E,\quad P_2=\frac{2}{3}\rho_E,\quad P_3=\rho_E,\quad \rho_E=\frac{E}{\Delta x^3}. \tag{8.34}$$

The minus sign in (8.33) and (8.34) has been omitted, it simply establishes whether an internal or external pressure expands or shrinks *V*. These results, which hold for photons or gases because no specific hypothesis has been made, are well known: P_1 holds for a light beam completely absorbed by the internal surface of *V*, whereas P_2 when waves or corpuscles bounce elastically; P_3 yields the well known law $P_3\delta V=\delta E$. As P_3 is due itself to elastic shocks of corpuscles against the internal walls of *V*, then $E=2E'/3$; *i.e.* *E* should be 2/3 of another energy E' that yields $PV=(3/2)E'$; is evident the connection of this last conclusion with the elementary kinetic theory of gases, where E' is easily demonstrable to be the average kinetic energy of molecules.

Skipping further considerations on this well known topic, return now to the Equations (8.25) and (8.26) to specify the result (8.34) in order to explain σ in the equation $e'=\sigma e$. The comparison of (8.32) and (8.26) suggests the correspondences

$$E\to\varepsilon,\quad \delta\Delta x\to\delta x,\quad \langle F\rangle\delta x=\varepsilon, \tag{8.35}$$

where the third position is the usual definition of force. So, owing to (8.26) and (8.18),

$$\frac{\delta E}{\delta\Delta x}\to\frac{\delta\varepsilon}{\delta x}=\frac{1}{\alpha}\frac{e'^2/\delta x}{\delta\Delta x},\quad \frac{E}{\Delta x}=\frac{\varepsilon}{\Delta x}=\frac{1}{\alpha}\frac{e^2}{\delta x^2},\quad \Delta x=\frac{n_v}{3}\delta x;$$

the third equation makes (8.33) compliant with (8.35). So the Equation (8.32) reads

$$\frac{e'^2}{\Delta x}=\frac{n_v}{3}\frac{e^2}{\delta x}, \tag{8.36}$$

whence, recalling the second (8.26), *i.e.*

$$e'=\pm e,\quad e'=\pm\frac{2}{3}e,\quad e'=\pm\frac{1}{3}e. \tag{8.37}$$

The second and third charges are consistent with quark charges, all with both signs correspondingly to $\pm e$, in the nuclear volume *V*. It is amazing the fact that even the quark charges appear here as a consequence of the dual wave/corpuscle behavior of matter and light, whereas their fractional character is reminiscent of the radiation/matter wave pressure in the volume enclosing them.

As a final remark, note that all charges can take both signs because *e* can be found in negative and positive energy states, as previously shown.

Now let us return to (8.24), to exemplify in a simple case how it is in fact calculable.

Let ε be the energy of a charged particle; (8.24) provides in principle the energy radiated per unit time with the help of (2.28)

$$\varepsilon\omega=\frac{F_\omega}{\hbar}\delta x\delta\varepsilon=F_\omega\frac{\delta\varepsilon}{\delta p}=F_\omega v.$$

It is easy to find the total power radiated $E/\delta t$ by such an accelerated particle, knowing that its charge is e and the change rate of its momentum is $\delta\dot{p}$. To solve this problem, however, more information is necessary about the link between radiation pressure and energy density. Once having found

$$P=\frac{2}{3}\frac{\varepsilon}{V}=\frac{2}{3}\frac{\varepsilon}{\left(c/\nu\right)^3},$$

being $\nu=1/\delta t$ the wave frequency, the non relativistic result as a function of the acceleration $a=\delta\dot{p}/m$ is obtained after having multiplied both sides by $\delta x^3/\delta t$ via the following chain of equations

$$\begin{aligned}P\frac{\delta x^3}{\delta t}&=\frac{2}{3}\frac{\varepsilon}{\delta t}\frac{\nu^3\delta x^3}{c^3}=\frac{2}{3}\frac{e^2}{\delta x\delta t}\frac{\nu^3\delta x^3}{c^3}=\frac{2}{3}\frac{e^2}{c^3}\frac{\left(\nu\delta x\right)^2}{\delta t^2}\\&=\frac{2}{3}\frac{e^2}{c^3}\frac{\delta v^2}{\delta t^2}=\frac{2}{3}\frac{e^2}{c^3}\frac{\left(\delta p/m\right)^2}{\delta t^2}=\frac{2}{3}\frac{e^2}{c^3}\left(\frac{\delta\dot{p}}{m}\right)^2.\end{aligned}\tag{8.38}$$

8.6. Short Range Nuclear Interactions

The Equations (3.12) and (8.4) suggest specific orders of magnitude significant to introduce short range forces. Skipping all theoretical details outside the purposes of the present paper, a few short comments are exposed below to highlight at least the essential features of these forces implementing only results so far obtained. Consider thus the following chances introduced by the general Equation (8.1)

$$F=\frac{\delta p}{\delta t}=v\frac{\delta p}{\delta x}=\frac{v}{c}\frac{\delta\left(pc\right)}{\delta x}$$

to infer, in agreement with (3.14) and the reasons therein explained,

$$F_s=\frac{\delta\left(pc\right)}{c\delta t},\quad F_w=\frac{v\delta\left(pc\right)}{c\delta x}\tag{8.39}$$

only in the first case δx has been replaced by $v\delta t$. The question that rises now is whether these expressions are mere equivalent ways of rewriting the same Equation (8.19) or they represent actually different force laws.

Preliminary inspection evidences that in both cases the force is defined via $\delta\varepsilon/length$ but in two different ways: F_w depends explicitly on v and δx, whereas F_s on $\delta x_s=c\delta t$ only. In the former case the interaction force is inversely proportional to δx; in the latter case the interaction energy $F_s\delta x_s$ increases with the distance δx_s, whatever this latter might be. This suggests the concept of "asymptotic freedom". Moreover the messenger particles of F_w carrying the interaction through $\delta x_w=v\delta t$ should be massive, owing to v; the messenger particles of F_s should be massless, as the force carried through δx_s involves c only. Consequently, one expects that for assigned δt the characteristic range measured by δx_s is greater than δx_w.

Taking according to (8.4) $\delta t\sim\delta x/c$ and $\delta\varepsilon\lesssim\hbar c/\delta x$, order of magnitude estimates of space ranges and time ranges that characterize F_s and F_w can be calculated utilizing the values (3.12). Particularly interesting is in this respect the range

$$\delta x_n = r_e - r_N = r_B \alpha^2 (1-\alpha).$$

1) Consider first F_s assuming preliminarily that the concerned interactions occur at the nuclear or sub-nuclear scale: *i.e.* reasonably δx_n concerns the interaction between different nucleons and in the nucleons themselves. Accordingly define the length

$$\sqrt{r_N r_e} = r_B \alpha^{2.5},$$

which is clearly an average value within δx_n, and introduce two complementary subranges δx_{s1} and δx_{s2} of the whole δx_n as follows

$$\delta x_{s1} = \sqrt{r_N r_e} - r_N = r_B \left(\alpha^{2.5} - \alpha^3\right), \quad \delta x_{s2} = r_e - \sqrt{r_N r_e} = r_B \left(\alpha^2 - \alpha^{2.5}\right). \quad (8.40)$$

It yields

$$\sqrt{r_N r_e} \approx 2.4 \times 10^{-14} \text{ cm}, \quad \lambda_C^{nucl} = 2.1 \times 10^{-14} \text{ cm};$$

moreover both subranges expressible through electron and nuclear range sizes yield

$$\begin{aligned} \delta x_{s1} &= r_B \left(\alpha^{2.5} - \alpha^3\right) \approx 2.2 \times 10^{-14} \text{ cm}, \\ \delta x_{s2} &= r_B \left(\alpha^2 - \alpha^{2.5}\right) \approx 2.6 \times 10^{-13} \text{ cm}. \end{aligned} \quad (8.41)$$

Is relevant the fact that that δ_{s1} is surprisingly close to the nucleon Compton lengths λ_C^{nucl} of both proton and neutron, which have in effect a similar order of magnitude. Also $\sqrt{r_N r_e}$ does so, which means that the nucleon mass represents the boundary value discriminating the interaction lengths δx_{s1} and δx_{s2} inside and outside the respective nucleon; this also explains the order of magnitude of the nucleon mass, indeed $\hbar c/\delta x_{s1} = 1.6 \times 10^{-24}$ g differs from the experimental nucleon mass by about 5.5% only. Here we take advantage of the fact that proton and neutron masses differ by less that 0.14% only. Hence the forces (8.4) defined by these ranges could concern both nucleons and their mutual interactions at distances consistent with the inequality (8.3); indeed, as expected, $\hbar c/\delta x_{s1} = \delta\varepsilon = 1.6 \times 10^{-3} \text{ erg} = 1 \text{ GeV}$ is related just to the order of magnitude of the nucleon mass, whereas $\hbar c/\delta x_{s2} = \delta\varepsilon_N = 1.2 \times 10^{-4} \text{ erg} = 0.08 \text{ GeV}$ is related to the binding energy between nucleons. In fact F_s is attractive, in agreement with the concept of "asymptotic freedom" already emphasized for quarks in nucleons less than δx_{s1} apart. The characteristic times are $\tau_1 \approx 7 \times 10^{-25}$ s and $\tau_2 \approx 9 \times 10^{-24}$ s. At this point it is possible a rough estimate of the stability of the nucleus comparing this energy $\delta\varepsilon_N$ with the Coulomb repulsion energy calculated via (8.23) approximately as $\varepsilon_{rep} \approx e^2/\delta x_{s2}$ between two protons δx_{s2} apart; $\varepsilon_{rep} \approx 9 \times 10^{-7}$ erg is negligible with respect to the attractive field in $\delta\varepsilon_N$. The fact that $\delta x_{s1} \gtrsim \lambda_C$ suggests that F_s should concern sub-nuclear particles that form protons and neutrons, which therefore are not elementary particles themselves. Hence the whole charge of proton and the null charge of the neutron can be due to nothing else but appropriate combinations of the e' fractional charges (8.37).

These considerations, well known and here shortly sketched only, are enough to conclude that F_s concerns the strong force. Further considerations are clearly outside the scopes of the present paper, merely aimed to show how to identify the fingerprints of the short range forces (8.39) in the conceptual frame hitherto outlined. Some more details are reported in [20].

2) Consider now F_w, noting that with the help of (2.28) the second Equation (8.39) reads also

$$F_w = \frac{v\delta(pc)}{c\delta x} = \frac{\delta\varepsilon}{\delta x} = \frac{\hbar}{\delta x\delta t}. \tag{8.42}$$

Being $v < c$ one expects $\delta x_w < c\delta t$; in other words, once having fixed δt, a shorter interaction rang $\delta x_w < \delta x_s$ is to be expected for δx_w of F_w. According to (8.4), $\delta x\delta\varepsilon \lesssim \hbar c$ implies information about δx_w compatible with $\delta\varepsilon_w$ in this case. To estimate v in the same reference system of (8.23), implement (8.23) supposing that an appropriate δx fulfills the condition

$$F = \frac{\hbar}{\delta x\delta t} \approx \frac{e^2}{\delta x^2\sqrt{1-v^2/c^2}}. \tag{8.43}$$

Elementary manipulations show that this position yields $\hbar v\sqrt{1-v^2/c^2} = e^2$ *i.e.* $v/c\sqrt{1-v^2/c^2} = \alpha$, whence the solutions $v/c \lesssim \alpha$ and $v/c \lesssim 1$. Hence, taking the same value of δt of (8.41) by comparison purposes, one finds two possible corresponding ranges

$$\delta x_\alpha \lesssim 7.7\times10^{-17}\ \text{cm},\quad \delta x_c \lesssim 1\times10^{-14}\ \text{cm},\quad \delta t_w \approx 3.5\times10^{-25}\ \text{s}. \tag{8.44}$$

In effect, extending the Equations (3.12) to the fourth power of α one finds the further length $\delta x_w = r_B\alpha^4 \approx 1.5\times10^{-17}$ cm necessary to include δx_α in the whole range of F_w. The first and second results are acceptable, as both yield space ranges shorter than that of both (8.41); the first value, in particular, yields according to the fourth (8.4)

$$\delta\varepsilon_w \lesssim \frac{\hbar c}{\delta x_\alpha} \approx 0.4\ \text{erg} = 255\ \text{GeV} \tag{8.45}$$

The existence of two range sizes (8.44) of δx compatible with (8.43) suggests that F_w should imply two different kinds of massive force carriers, reasonably with and without charges; if so, then the charges must have opposite signs. This kind of interaction needs thus three kinds of carriers. Assuming charged and neutral carriers of masses $m_\pm$ and m_0, just a few considerations are enough to infer significant information on the masses of these messengers.

Implementing this assumption to establish the energy balance governing the formation of the carriers, the results are in full agreement with the experimental data.

Is reasonable the idea of regarding the cluster of messengers as a system of particles interacting themselves in order that the gain of binding energy of the charges accounts not only for their own masses m_+ and m_- but also for that of m_0. Is interesting the energy balance of the charged carriers according to the electromagnetic Equation (8.43). Start with the energy of a Coulomb system with a nucleus of mass $m_\pm$ formed by either charge, e.g. m_+, in the field of which interacts the other charge, e.g. m_-, at average distance r_B; the subscript "*B*" stands for "bound", whereas the simplest hypothesis on the masses is $m_+ = m_-$. Consider preliminarily that such a system can be described as shown in subsection 3.2: the reasoning introduced to describe the electron charge around the nuclear charge holds in principle also for integer spin charged particles. The early hydrogenlike atom was introduced before the concept of spin, which became essential to account for the electron pile up in many electron atoms [7] according to the exclusion principle and for the possible presence of an external field. The success of Bohr's idea was allowed by the fact that the spin-orbit and spin-spin interaction between electron and nucleus are both small with respect to the Coulomb interaction. Consider at this point uniquely (3.1) that has general validity and skips, as shown in section 3.3, the operator formalism implementing wave functions along with all related implications: e.g. it is known that a 0 spin particle requires a 4 dimensional scalar wave function, whereas a spin 1 particles requires a three component wave function. On the one hand (3.1) has 4 dimensional character as it merges space and time coordinates through the respective uncertainty ranges, to which are related energy and momentum ranges too. On the other hand the necessity of describing the particle in any reference system is in fact ensured by (3.1) according to (3.2). Implement thus the electromagnetic interaction only to describe via (3.1) even a system of spin 1 charged bosons m_+ and m_- trusting that the steps from (3.3) to (3.9) still hold at least approximately also now; the comparison with the experimental data will be the decisive benchmark to assess the validity of these considerations. So it is possible to write for the system of boson charges $\varepsilon_n = e^4 m_r / 2n^2\hbar^2$, see Equation (3.7) with $Z = 1$, where m_r is the reduced mass of the concerned system. Accordingly (3.9) yields

$$\varepsilon_B = -\frac{1}{2}\frac{\alpha^2}{n^2} m_r c^2 = -\frac{1}{2}\frac{\hbar c\alpha}{r_B}, \quad r_B = \frac{n^2\hbar c}{\alpha m_r c^2}, \quad m_r = m_\pm/2 \tag{8.46}$$

for charges of equal mass. Putting $n = 1$ and including α into m_r, this equation reads

$$\varepsilon_B = -\frac{1}{2} m_r' c^2 = -\frac{1}{2}\frac{\hbar c}{r_B'}, \quad m_r' = m_r\alpha^2, \quad r_B' = \frac{r_B}{\alpha} = \frac{\hbar}{m_r' c}; \tag{8.47}$$

with these positions ε_B depends explicitly on r_B' and m_r' only, no longer on r_B and m_r. Is of interest now an appropriate r_B' compliant with ε_B and such that

$$(m_0 + 2m_\pm)c^2 = \frac{1}{2}\frac{\hbar c}{r_B'} = \frac{\pi\hbar c}{n'\lambda}, \quad 2\pi r_B' = n'\lambda \tag{8.48}$$

with n' integer, the second equation is a well known condition of the wave mechanics already implemented in (3.8) with the same physical meaning. Let the shortest wavelength λ be the Compton length of either $m_\pm$, as suggested by the Equation (3.6) and [7], *i.e.*

$$\lambda = \frac{\hbar}{m_\pm c};$$

then, replacing $m_\pm c^2 = \hbar c/\lambda$ into (8.48) one finds

$$m_0 c^2 = (\pi - 2n')\frac{\hbar c}{n'\lambda}.$$

In conclusion, comparing with (8.48) and putting $n' = 1$, it is possible to write

$$\frac{m_0 c^2}{(m_0 + 2m_\pm)c^2} = \frac{\pi - 2}{\pi}, \quad \frac{m_\pm c^2}{(m_0 + 2m_\pm)c^2} = \frac{1}{\pi}.$$

These results are verifiable by comparison with the experimental masses $m_\pm$ and m_0:

$$m_0 = 91.19\ \text{GeV}, \quad m_\pm = 80.39\ \text{GeV},$$

$$\frac{m_0 c^2}{(m_0 + 2m_\pm)c^2} = 0.36, \quad \frac{m_\pm c^2}{(m_0 + 2m_\pm)c^2} = 0.32,$$

which in effect compare well with $(\pi - 2)/\pi = 0.36$ and $1/\pi = 0.32$ respectively.

Moreover note that (8.47) regards by definition r'_B as average distance between m_+ and m_-, whereas δx_a of (8.44) is by definition the total range of F_w; therefore one infers that reasonably $r'_B = \delta x_a/2$. Hence, according to (8.48), $\varepsilon_B = -\hbar c/\delta x_a$ is the binding energy gain available to create the masses $(m_0 + 2m_\pm)c^2$. In effect one finds that the total energy ε_w related to F_w is

$$\varepsilon_w = -\varepsilon_B = \frac{\hbar c}{\delta x_\alpha} = 255\ \text{GeV} = (m_0 + m_+ + m_-)c^2, \tag{8.49}$$

in agreement with (8.45) and with the experimental masses. The Equations (8.47) and (8.46) differ in fact only formally; once having removed α merely including it in the reduced mass of the system as a numerical scale factor, someway analogous to (3.12), it appears that F_w is different from but closely related to the electromagnetic interaction constant via the linked energy scale factor.

In effect, considering m'_r and r'_B, and not m_r and r_B, one calculates experimental masses of the force carriers and reasonable estimate of the interaction range $\delta x_\alpha \approx 7.3\times10^{-17}$ cm that agree with the total energy (8.45); these values support the idea that a hydrogenlike system bound by electromagnetic interaction via photon carriers turns into a short range interaction system via massive carriers. There appears in this way the link between electromagnetic and weak interactions.

At this point, something else about δx_α can be still inferred to confirm that F_w corresponds to the weak interaction. Helps in this respect the first (8.4) $FV = \hbar\delta x^2/\delta t$ that reads

$$\varepsilon_w V = \frac{\hbar}{\delta t}\delta x^3, \quad \varepsilon_w = \frac{\hbar c}{\delta x_\alpha}$$

whence, with the help of the values (8.44),

$$\varepsilon_w V = \frac{\hbar c}{\delta x_\alpha}\delta x_\alpha^3 = 2\times 10^{-49}\ \text{erg}\cdot\text{cm}^3; \tag{8.50}$$

this is the Fermi constant characterizing the weak interactions.

This result is more than mere fingerprint of weak interactions; interesting information can be inferred from it expressing appropriately energy and volume inherent this result. Write

$$2\times 10^{-49} = \varepsilon_F \lambda_F^3 = \varepsilon_F\left(\frac{\hbar}{m_F c}\right)^3 = \varepsilon_F\left(\frac{\hbar c}{\varepsilon_F}\right)^3, \quad \varepsilon_F = m_F c^2 \tag{8.51}$$

the energy ε_F that defines the characteristic Fermi constant has been expressed via Compton length of the characteristic mass m_F that in turn defines ε_F too. Hence

$$\varepsilon_F = \sqrt{\frac{(\hbar c)^3}{2\times 10^{-49}}} = 0.397\ \text{erg} = 255\ \text{GeV}. \tag{8.52}$$

It is not surprising that one finds once more the value of total energy of this kind of interaction. Implement now the idea that actually the energy (8.49) is degenerate: it consists of m_- moving in the field of m_+ or, identically, from m_+ moving in the field of m_-. As both configurations can coexist consistently with the unique reduced mass m_r (8.46) that calculates e_B of (8.47), it is reasonable to regard the value (8.49) as the sum of both allowed chances; this means that the total energy refers to the total volume calculated via (8.51), so that each configuration has energy

$$\varepsilon_H = \frac{1}{2}\varepsilon_F = 0.2\ \text{erg} = 127\ \text{GeV}. \tag{8.53}$$

Note that this value is also consistent with that inferred through a characteristic range similarly as done in (8.40) and (8.41)

$$\begin{aligned}&\delta x_H = r_B\left(\alpha^{3.5} - \alpha^4\right) = 1.6\times 10^{-16}\ \text{cm},\\ &\frac{\hbar c}{\delta x_H} = 0.197\ \text{erg} = 123\ \text{GeV}.\end{aligned} \tag{8.54}$$

8.7. The Dirac and Lamb Equations

This section generalizes the results of the Section 3.2 obtained implementing the non-relativistic equation $p^2 = p_r^2 + M^2/2r^2$. The following considerations show how to describe a relativistic hydrogenlike system replacing the classical position (3.4) with the series expansion

$$\varepsilon = \ell_0 \delta\dot{p} = \sigma_0 + \frac{\sigma_1}{\Delta r} + \frac{\sigma_2}{\Delta r^2} + \frac{\sigma_3}{\Delta r^3} + \cdots \tag{8.55}$$

expressing $\delta\dot{p} = \delta\dot{p}\left(\Delta r^{-1}\right)$ as $\delta\dot{p} = \sum_{i=0}\sigma_i \Delta r^{-i}$ similarly as done in the Equations (8.5) or (8.6). In effect even the term $\Delta p_r^2/2m$ of (3.4) can be written as $\sigma_1'/\Delta r^2$ with coefficient $\sigma_1' = (n\hbar)^2/2m$; so (3.4) is actually a particular case of the series (8.55) truncated at the second order. Of course ℓ_0 is an arbitrary constant length that introduces the energy corresponding to force $\delta\dot{p}$. Although $\delta\dot{p}$ vanishes at the infinity, the arbitrary constant $\sigma_0 = mc^2$ accounts for the electron rest mass energy. Hence it is possible to write

$$\begin{gathered}\delta\varepsilon = \varepsilon_2 - \varepsilon_1 = \ell_0\delta\dot{p} - mc^2 = \frac{\sigma_1}{\Delta r} + \frac{\sigma_2}{\Delta r^2} + \frac{\sigma_3}{\Delta r^3} + \cdots, \\ \Delta r = r_2 - r_1, \quad \Delta p_r = p_{r2} - p_{r1},\end{gathered} \tag{8.56}$$

being Δp_r the radial momentum range conjugate to Δr. All range boundaries are of course arbitrary. In this way we deliberately waive introducing explicitly radial and angular momenta exploited in section 3.2, but implement directly the fundamental Equation (3.1). Multiply both sides of (8.56) by mc^2 so that

$$\left(\ell_0\delta\dot{p} - \frac{\sigma_1}{\Delta r}\right)mc^2 - \left(mc^2\right)^2 = \frac{\left(\Delta p_r c\right)^2 m\sigma_2}{\left(n\hbar\right)^2} + \frac{\left(\Delta p_r c\right)^3 m\sigma_3}{\left(n\hbar\right)^3 c} + \cdots,$$

The series truncated at the third order yields

$$\begin{gathered}\delta\left(\varepsilon^2\right) = \varepsilon_2^2 - \varepsilon_1^2 = \frac{\left(\Delta p_r c\right)^2 m\sigma_2}{\left(n\hbar\right)^2} + \left(mc^2\right)^2 + \frac{\left(\Delta p_r c\right)^3 m\sigma_3}{\left(n\hbar\right)^3 c}, \\ \delta\left(\varepsilon^2\right) = mc^2\ell_0\delta\dot{p} - mc^2\frac{\sigma_1}{\Delta r}.\end{gathered}$$

As the coefficient σ_2 has not yet been defined, it is convenient to turn this equation into

$$\delta\left(\varepsilon^2\right) = \left(\Delta p_r c\right)^2 + \left(mc^2\right)^2 + \frac{\left(\Delta p_r c\right)^3 m\sigma_3}{\left(n\hbar\right)^3 c} + \cdots, \quad \sigma_2 = \frac{\left(n\hbar\right)^2}{m};$$

then, dividing both sides by $\left(mc^2\right)^2$, one finds

$$\begin{gathered}\frac{\varepsilon_2^2}{\left(mc^2\right)^2} = \left(\frac{\Delta p_r c}{mc^2}\right)^2 + 1 + \Theta, \\ \Theta = \frac{\varepsilon_1^2}{\left(mc^2\right)^2} + \frac{\left(\Delta p_r c\right)^3 m\sigma_3}{\left(mc^2\right)^2\left(n\hbar\right)^3 c} = \frac{\varepsilon_1^2}{\left(mc^2\right)^2} + \frac{\sigma_3}{mc^2\Delta r^3}.\end{gathered} \tag{8.57}$$

Next write

$$\frac{\Delta p_r c}{mc^2} = \frac{p_{r2}c}{mc^2} - \frac{p_{r1}c}{mc^2} = si\sqrt{\left(\frac{\varepsilon_2^2}{\left(mc^2\right)^2} - \Theta\right) - 1}, \quad si = \pm 1 \tag{8.58}$$

so that, subtracting ε_2/mc^2 at both sides, one finds

$$\frac{p_{r2}c - \varepsilon_2}{mc^2} = b + si\sqrt{\left(a^2 - \Theta\right) - 1} - a, \quad b = \frac{p_{r1}}{mc^2}, \quad a = \frac{\varepsilon_2}{mc^2} \tag{8.59}$$

and then

$$\frac{\left(mc^2\right)^2}{\left(p_{r2}c - \varepsilon_2\right)^2} = \left(b + si\sqrt{\left(a^2 - \Theta\right) - 1} - a\right)^{-2}.$$

According to the first (8.57) $\varepsilon_2^2-(1+\Theta)(mc^2)^2=(\Delta p_r c)^2$; thus the last equation reads

$$\frac{\varepsilon_2^2}{(1+\Theta)(p_{r2}c-\varepsilon_2)^2}-\frac{(\Delta p_r c)^2}{(1+\Theta)(p_{r2}c-\varepsilon_2)^2}=\left(b+si\sqrt{(a^2-\Theta)-1}-a\right)^{-2}. \tag{8.60}$$

Put now

$$p_{r2}c-\varepsilon_2=\pm\Delta p_r c. \tag{8.61}$$

This position has two implications: the first replacing it in (8.57)

$$\varepsilon_2^2-(1+\Theta)(mc^2)^2=(p_{r2}c-\varepsilon_2)^2 \tag{8.62}$$

and the second replacing in (8.60)

$$\frac{\varepsilon_2^2}{(1+\Theta)(p_{r2}c-\varepsilon_2)^2}=\left(b+si\sqrt{(a^2-\Theta)-1}-a\right)^{-2}+\frac{1}{1+\Theta}. \tag{8.63}$$

Note that we have introduced four conditions: *a* and *b* in (8.59) plus (8.62) and (8.63); the unknowns in these equation are $p_{r1},p_{r2},\varepsilon_2,\Theta$. In principle the system appears solvable.

Taking the reciprocal of both sides one finds

$$si'\left(\frac{p_{r2}c}{\varepsilon_2}-1\right)=\frac{1}{\sqrt{1+\Theta}}\left[\left(b+si\sqrt{(a^2-\Theta)-1}-a\right)^{-2}+\frac{1}{1+\Theta}\right]^{-1/2},\quad si'=\pm1. \tag{8.64}$$

The notations si and si' have been introduced to allow that the upper and lower signs in (8.58) and (8.56) are independent each other. Then it is possible that

$$-\frac{p_{r2}c}{\varepsilon_2}=\frac{1}{\sqrt{1+\Theta}}\left[\left(b+si\sqrt{(a^2-\Theta)-1}-a\right)^{-2}+\frac{1}{1+\Theta}\right]^{-1/2}-1,\quad si'=-1 \tag{8.65}$$

or

$$\frac{p_{r2}c}{\varepsilon_2}=\frac{1}{\sqrt{1+\Theta}}\left[\left(b+si\sqrt{(a^2-\Theta)-1}-a\right)^{-2}+\frac{1}{1+\Theta}\right]^{-1/2}+1,\quad si'=1 \tag{8.66}$$

Subtracting (8.65) from (8.66) with the same sign si, one finds $(p_{r2}c/\varepsilon_2)_+-(p_{r2}c/\varepsilon_2)_-=2$: this suggests that the left hand side of these equations must have the form E/mc^2, so that

$$E_+-E_-=2mc^2.$$

The minus sign at the left hand side of (8.65) represents binding energy of the electron to the nucleus in mc^2 units; the first addend at the right hand side represents the energy gain with respect to that of the free electron in either energy state.

The previous algebraic steps aimed just to find an equation introducing the ratio $p_{r2}c/\varepsilon_2$. In effect this ratio is significant because, according to (8.61), if $p_{r2}c\to\varepsilon_2$ then $\Delta p_r\to0$ and thus $\Delta r\to\infty$ whereas p_r becomes constant; this is the limit case of free electron. Indeed $\Delta p_r\neq0$ implies binding energy, since the electron takes random values of radial momentum between p_{r1} and

p_{r2} depending on its finite random distance from the nucleus.

Examine the result (8.65) putting first $\Theta = 0$: in fact, according to the second (8.57), this occurs putting $\sigma_3 = 0$ and in (8.55) and $\varepsilon_1 = 0$, *i.e.* considering the energy ε_2 only instead of the energy range $\delta\varepsilon$ (Heisenberg compliant quantum case). Now require that

$$\frac{E_-}{mc^2} = \left[\left(b + si\sqrt{a^2 - 1} - a\right)^{-2} + 1\right]^{-1/2} - 1 \qquad (8.67)$$

must be compatible with the non-relativistic quantum Equations (3.5) and (3.9); in other words, in (8.67) must somehow appear not only αZ but also n and $l(l+1)$ as well. To fulfill this boundary condition as a limit case for small values of αZ, put in (8.67))

$$a = \frac{a_o}{\alpha Z}, \quad b = \frac{b_o}{\alpha Z}; \qquad (8.68)$$

in effect, replacing and expanding in series around $\alpha Z = 0$, (8.67) becomes

$$\left.\frac{E_-}{mc^2}\right|_{\alpha Z \to 0} = \frac{(\alpha Z)^2}{2\left(b_o + sia_o - a_o\right)^2}.$$

Considering in particular $si = 1$, the boundary condition requires $b_o = n$. Once having identified b_o, regard then a_o and a_o^2 in order to be compliant with l and $l(l+1)$ of Equation (3.5), while also fulfilling (3.9) and (3.10). This suggests reasonably $a_o = l + 1/2 \pm s$, being s the electron spin; in effect, depending on the sign, a_o becomes l or $l+1$. In conclusion

$$\frac{E_-}{mc^2} = \left[\left(\frac{\alpha Z}{n + \sqrt{(j+1/2)^2 - (\alpha Z)^2} - (j+1/2)}\right)^2 + 1\right]^{-1/2} - 1, \quad j = l \pm s, \quad \Theta = 0 \qquad (8.69)$$

This is the Dirac equation, which however becomes in the present approach particular case of an even more general equation including Θ as well. The Equation (8.65) reads indeed

$$\frac{E_-}{mc^2} = \frac{1}{\sqrt{1+\Theta}}\left[\left(\frac{\alpha Z}{n + \sqrt{(j+1/2)^2 - (1+\Theta)(\alpha Z)^2} - (j+1/2)}\right)^2 + \frac{1}{1+\Theta}\right]^{-1/2} - 1$$

which removes the degeneracy of states with equal n and j of the Dirac equation and also suggests that a further physical effect related to Θ not concerned in (8.69) is still hidden in this result. In effect the Dirac equation becomes in this approach the zero order approximation of a more complex energy function whose series expansion reads

$$\frac{E_-}{mc^2} = E_{Dir} + \left.\frac{\partial E_-}{\partial \Theta}\right|_{\Theta=0} \Theta + \cdots \qquad (8.70)$$

It is evident that (8.70) removes the degeneracy of the $2p_{1/2}$ and $2s_{1/2}$ states: indeed, whatever the actual analytical form of Θ might be, calculating the energy difference of these states one finds

$$\left.\frac{E_-}{mc^2}\right|_{2p_{1/2}}-\left.\frac{E_-}{mc^2}\right|_{2s_{1/2}}=\left(\left.\frac{\partial E_-}{\partial\Theta}\right|_{\Theta=0}\Theta\right)_{2p_{1/2}}-\left(\left.\frac{\partial E_-}{\partial\Theta}\right|_{\Theta=0}\Theta\right)_{2s_{1/2}}\neq 0.$$

Owing to the physical dimensions $energy\times length^3$ of the coefficient σ_3 in the second (8.57), it is easy to guess the order of magnitude of the second addend of Θ according to the following reasonable positions

$$\sigma_3\approx\varepsilon_{bohr}\times\lambda_C^3,\quad \Delta r\approx r_{bohr},\quad \frac{\sigma_3}{\Delta r^3}\approx\frac{\varepsilon_{bohr}\lambda_C^3}{r_{bohr}^3},$$

being λ_C the electron Compton length. So it follows from the second (8.57) with the help of (3.4) and (3.8)

$$\frac{\varepsilon_{bohr}\lambda_C^3}{mc^2r_{bohr}^3}=-\frac{1}{2}\left(\frac{\alpha Z}{n}\right)^5,\quad \Theta\approx\frac{\varepsilon_1^2}{\left(mc^2\right)^2}-\frac{1}{2}\left(\frac{\alpha Z}{n}\right)^5. \tag{8.71}$$

The second addend of Θ in (8.71) is the signature of the radiative energy displacement due to the interaction of the electron with the quantum vacuum, known as electron driven vacuum polarization effect; the analytical form of the first addend represented by ε_1, not yet concerned explicitly, is at present still under investigation. So, even without detailed calculations in this respect, appear two relevant facts: 1) the first three terms of (8.55) are enough to infer the Dirac equation; 2) the cubic term and the implementation of the energy range $\delta\varepsilon$ instead of a unique energy term ε are essential to infer contextually the Lamb energy shift too.

9. Discussion

The present model has concerned several topics of fundamental physics self-consistently inferred uniquely from the concept of evolution inherent the definitions (1.11) and (1.12). The concepts of mass, momentum, energy and electric charge, obviously missing in these equations, have been uniquely and self-consistently introduced through the fundamental constants of Nature. It appears also significant the chance of describing the Universe according to laws inferred from the change of a unique primordial function ψ, even regardless of a specific and detailed knowledge about the function that is changing itself: it is instead crucial how it changes.

The fact of having introduced an initial function and its actual space time evolution, has been proven enough to infer contextually quantum uncertainty and relativistic results in a surprisingly straightforward way even regardless of any deterministic metric and without hypotheses "ad hoc".

On the one side the necessity of quantized physical laws is implied by the concept itself of uncertainty, Equations (3.1) and (3.2), on the other side special and general relativity are implied by the space time frame under the condition of its holistic evolution. The reverse reasoning is also true: the foundation of quantum and relativistic theories are the fingerprint of an evolving Universe, whose evolution is governed by a few constants in which are nested the essential dynamical variables of interest for the everyday experimental activity, mass, energy charges and so on.

The modern physics is essentially wave physics. This is because the Bohr atom first opened the way to the hydrogenlike atoms and thus to the probabilistic interpretation of the wave functions. Next Schrödinger further enhanced this conceptual path including in the wave function the potential term and thus the electron correlation in many electron atoms and ions. Eventually a further step ahead was accomplished by Dirac: with its relativistic hydrogen atom, He has in fact introduced the quantum field theory. Yet all these physical models implemented wave formalism. The present paper, instead, introduces and contextually exploits the corpuscle nature of the particles constituting the matter, appropriately integrated with their wave nature when necessary. The subsection 3.2 has been reported just to clarify this point. In this way is irrelevant the theoretical problem raised by many physicists about why $\Psi^*\Psi$, and not Ψ itself, has physical meaning [21]; moreover the approach to the various equations of quantum and relativistic physics appears not only simpler but also the equations themselves are more interconnected. The Heisenberg principle has negative content; the statistical formulation of the space time uncertainty has instead a highly positive content as it shares both quantum and relativistic theories. Regarding a fundamental statement the uncertainty and following the approach shortly sketched in Section 3.2 the EPR paradox would be meaningless because the concept of distance is missing; the uncertainty ranges waive since the beginning conceptually, and not as a sort of approximation useful to simplify calculations, the concepts of local space time coordinates necessary to define "superluminal" distances.

As concerns the quantum way of describing the reality, these basic concepts can be summarized as follows

$$\begin{array}{ccc} -\frac{\hbar}{i}\frac{\partial}{\partial x} & & \frac{\hbar}{i}\frac{\partial}{\partial t} \\ \text{operator} & & \text{operator} \\ \updownarrow & & \updownarrow \\ p & \text{dynamical variables} & \varepsilon \\ \updownarrow & & \updownarrow \\ \overbrace{\Delta p}^{\text{uncertainty range}} & & \overbrace{\Delta \varepsilon}^{\text{uncertainty range}} \\ \Delta x & \hbar, c, G, \alpha & \Delta t \end{array} \tag{9.1}$$

The upper part deals with differential equations that by definition describe the local properties of the solution of the pertinent wave equation; the lower part describes instead the system regardless of its local properties and thus without need of solving the pertinent differential equations. In principle both approaches are equivalent, although the operator formalism is a byproduct of the quantum uncertainty; in practice, however, the problem is to see which approach is more effective in describing the quantum properties of the Universe regardless of the local and deterministic tensor formalism. It is worth recalling that all papers based only on the Equation (3.1) only, allowed to obtain the most significant results of both general relativity and quantum physics [20]; in the latter case, in particular, the usual positions (3.28) introducing the operator formalism of wave mechanics according to (3.66) and (3.67) are systematically replaced by the uncertainty positions

$$x \to \delta x, \quad p \to \delta p, \quad t \to \delta t, \quad \varepsilon \to \delta\varepsilon, \tag{9.2}$$

while obtaining results identical to that of the standard wave formalism, as shortly shown in subsection 3.2.

The classical dynamical variables p and ε are to be regarded equivalently as quantum differential operators or quantum uncertainty ranges: this implies that actually it is necessary neither to solve the Schrödinger equation of wave mechanics nor the tensor calculus of relativity. The form (3.2) expressing the quantum uncertainty, apparently weird, shows the quantum equivalent of the relativistic covariance: the Equations (3.1) could seem defined in some particular reference system, instead (3.2) show that whenever the dynamical variables are replaced by the respective uncertainty ranges about which nothing is known in the sense highlighted in the subsection 3.2, the dependence of any formula on a particular reference system, inertial or not, is lost. So the independence of formulas on any particular R is ensured by (3.1), despite their different forms in R and R'; however holds the more substantial fact the any formula inferred from (3.1) has validity in any R': this is the profound reason why relativistic formulas can be inferred from (3.1).

Yet, the rational foundation of everything is just the conceptual impossibility of knowing everything.

10. Conclusions

As stated in Section 1, part of this paper aimed to find known results as a test of validity of the present theoretical model. Besides well known results, explicitly quoted throughout the exposition, the model has also provided original results:

- Evolutionary imprinting and derivation of physical laws.
- Possible granular structure of the space time.
- Possible quantization of the temperature.
- Lnk between entropy, phase space and space time.
- Link between Van der Waals equation and quantum zero point state of matter.
- Link between relativity and quantum gravity.
- Probabilistic link between corpuscle and wavelike behavior of matter.
- Link between operator and uncertainty driven approach to quantum problems.
- Generalization of Dirac equation to include the Lamb effect.

Moreover:

- The Equations (4.4) and (3.13) show that even a small mass m_0 can take large values of kinetic mass m for $v \to c$; also, (4.6) shows that just m is the classical mass.

- The Equation (6.19) has shown the existence of finite vacuum energy density η, to which corresponds according to (8.34) a pressure $P_{vac} = 2.2\times10^{-8}\ \text{dyn}/\text{cm}^2$.
- The Equation (4.7) has shown that the corpuscle/wave behavior of matter has probabilistic character and that this probability involves the ratios m_0/m and v_g/c.
- The velocity dependence of mass shows that $m \to m_0$ implies $v \to 0$, whatever m_0 might itself be; contextually, increasing m_0 to m' means decreasing v_g from c to a smaller value v'.
- Short notes, although necessarily incomplete, emphasize the essential fingerprints of the strong and weak interactions, Equations (8.37), and contextually also the gravity force and Maxwell equations.
- The model explains why un upper limit of velocity, *c*, must necessarily exist.

These short remarks are enough to conclude that the present model fits the basic concepts of thermodynamics and fundamental forces of nature merging concept of quantum and relativistic physics.

Conflicts of Interest

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

References

[1] Wigner, E. (1960) The Unreasonable Effectiveness of Mathematics in the Natural Sciences. Communications in Pure and Applied Mathematics, vol. 13, No. I (February). John Wiley and Sons, Inc., New York.

[2] Russell, B. (1907) The Study of Mathematics. The New Quarterly 1 (Nov) Repr. Philosophical Essays, Longmans, Green, and Co., 1910; Mysticism and Logic and Other Essays, London, Longmans, Green, and Co., 1918, pp. 58-73 Page numbers are to ML 1918.

[3] Zwiebach, B. (2004) A First Course in String Theory. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511841682

[4] Tosto, S. (1996) *Il Nuovo Cimento B*, **111**, 193-215. https://doi.org/10.1007/BF02724645

[5] Tosto, S. (1996) *Il Nuovo Cimento D*, **18**, 1363-1394. https://doi.org/10.1007/BF02453780

[6] Tosto, S. (2016) *Journal of Modern Physics*, **7**, 1668-1701. https://doi.org/10.4236/jmp.2016.713152

[7] Tosto, S. (2017) *Open Journal of Physical Chemistry*, **7**, 89-121. https://doi.org/10.4236/ojpc.2017.73007

[8] Amelino Camelia, G. (2002) *Physics Letters B*, **528**, 181-187. https://doi.org/10.1016/S0370-2693(02)01223-6

[9] Cahn, J.W. (1961) *Acta Metallurgica*, **9**, 795-801. https://doi.org/10.1016/0001-6160(61)90182-1

[10] Tosto, S. (2016) *Journal of Advances in Physics*, **11**, 3408. https://doi.org/10.24297/jap.v11i5.362

[11] Landau, L. and Lifchitz, E. (1966) Theorie du Champ. Ed. MIR, Moscow.

[12] Hevesy, G. and Obrutsheva, C. (1925) *Nature*, **115**, 674-675. https://doi.org/10.1038/115674a0

[13] Tosto, S. (2009) An Analysis of States in the Phase Space: The Specific Heat of Metals. Recent Research Developments in Physics, S.G. Pandalai Ed., Vol. 8, Trivandrum, India, p. 7.

[14] Landau, L. and Liftchitz, E. (1969) Mechanique. MIR Editions, Vol. 1, Moscow, p. 57.

[15] Tosto, S. (2014) *International Journal of Physics and Astronomy*, **27**, 1136-1156.

[16] Tosto, S. (2013) *American Journal of Space Science*, **1**, 22-32. https://doi.org/10.3844/ajssp.2013.22.32

[17] Tosto, S. (2013) *Physics International*, **4**, 135-151. https://doi.org/10.3844/pisp.2013.135.151

[18] Tosto, S. (2014) *International Journal of Physics and Astronomy*, **27**, 1136.

[19] https://en.wikipedia.org/wiki/Planck_units

[20] Tosto, S. (2015) An Analysis of the States in the Phase Space: From Quantum Mechanics to General Relativity. Arxiv gr-qc/0807.1011, Vol. 11, 149.

[21] Leonhardt, U. and Paul, H. (1995) *Progress in Quantum Electronics*, **19**, 89-130. https://doi.org/10.1016/0079-6727(94)00007-L

Scientific Research Publishing

Journal of Modern Physics, 2018, 9, 2594-2622
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Photoionization Study of Cl II, Ar II and Kr II Ions Using the Modified Atomic Orbital Theory

Abdourahmane Diallo[1], Mamadou Diouldé Ba[1], Jean Kouhissoré Badiane[1], Momar Talla Gning[1], Malick Sow[2], Ibrahima Sakho[1]*

[1]Department of Physics, UFR Sciences and Technologies, University Assane Seck of Ziguinchor, Ziguinchor, Senegal
[2]Department of Physics, Faculty of Sciences and Techniques, University Cheikh Anta Diop of Dakar, Dakar, Senegal
Email: *aminafatima_sakho@yahoo.fr

How to cite this paper: Diallo, A., Ba, M.D., Badiane, J.K., Gning, M.T., Sow, M. and Sakho, I. (2018) Photoionization Study of Cl II, Ar II and Kr II Ions Using the Modified Atomic Orbital Theory. *Journal of Modern Physics*, **9**, 2594-2622.
https://doi.org/10.4236/jmp.2018.914162

Received: September 30, 2018
Accepted: December 24, 2018
Published: December 27, 2018

Abstract

Resonance energies of the Cl II-$[3s^23p^3(^2D_{5/2})]n$d and $[3s^23p^3(^2P_{3/2})]n$d, Ar II-$3s^23p^4(^1D_2)n$s, nd and of the Kr II $[4s^24p^4(^1D_2)]n$s, nd and $4s^24p^4(^3P_2,^3P_1)]n$s, $4s^24p^4(^3D_2)]n$s, nd and $4s^24p^4(^3D_2,\ ^1S_0)]n$s, nd Rydberg series are reported. Natural widths of the Ar II-$[3s^23p^4(^1D_2)]n$s, nd series are also reported. Calculations are done in the framework of the Modified Atomic Orbital Theory (MAOT). Excellent agreements are obtained with available theoretical and experimental data. High lying accurate resonance energies up to $n = 40$ are tabulated. The possibility to use the MAOT formalism report rapidly with an excellent accuracy the position of the excitation resonances as well as their width within simple analytical formulae is demonstrated.

Keywords

Resonance Energies, Rydberg Series, Natural Widths, Modified Atomic Orbital Theory (MAOT)

1. Introduction

In many astrophysical systems such as stars and nebulae, the main process governing light-atomic species interaction is Photoionization. Of great important ions interesting to investigate are Cl II (Cl^+), Ar II (Ar^+) and Kr II (Kr^+) ions. As far as Cl II is concerned, its interest is connected with it abundance in photoionized astrophysical objects. In the past, various studies have indicated the great importance of Cl II ions abundances for understanding extragalactic HII regions [1]. In addition, emission lines of Cl II ions have been observed in the spectra of the Io torus [2] and in the optical spectra of planetary nebulae NGC 6741 and IC 5117 [3]. In a very recent past, Hernández *et al.*, [4] measured at the Advanced Light Source at Lawrence Berkeley National Laboratory absolute photoionization cross-sections for the of Cl II ions using the merged beams photon-ion technique at a photon energy resolution of 15 meV in the energy range 19 - 28 eV. Using the Dirac-Coulomb *R*-matrix (DCR) method, McLaughlin, [5] performed calculations in the same photon energy range that in the ALS experiments [4] to assign and identify the resonance series in the ALS spectra of the Cl II ions. In these experimental works, the $3s^23p^3nd$ states have been identified in the Cl II spectra as the prominent Rydberg series belonging to the 3p → *n*d transitions. The weaker $3s^23p^3ns$ Rydberg series, identified as 3p → *n*s transitions and window resonances $3s3p^4(^4P)$ *n*p features, due to 3s → *n*p transitions, have also been found in the spectra [5]. Besides, one of the important elements to study is Argon present in several astrophysical systems. An overabundance of the argon element in the spectra of X-rays of yellow supernovas was revealed by the satellite Chandra [6] dedicated to analyze the stellar object spectra. Furthermore, spectral rays of the argon element were observed in the optical spectra of planetary stars and nebulae [7] [8]. These few examples show the importance of the photoionization study of the argon element from the perspective of astrophysics. In a recent past, Covington *et al.*, [9] performed the first experimental measurements of the photoionization cross-section of the Ar II ion. These authors also determined the resonance energies and natural widths related to the dominant Rydberg series $3s^23p^4(^1D_2)ns$, *n*d and $3s^23p^4(^1S_0)ns$, *n*d in the emission spectra of the Ar II ions. These energies were relatively measured at the metastable state Ar II ($^2P°_{1/2}$) and at the ground state Ar II ($^2P°_{3/2}$). As far as Krypton is concerned, it is also an element of major importance for diagnosing stellar plasmas such as stars and planetary nebulae as well as for diagnosing laboratory plasmas such as those obtained by inertial fusion. In a recent past, Hinojosa *et al.*, [10] experimentally studied the photoionization of the Kr II ion at ALS at Berkley in the photonic energy range of 23 - 39 eV. In the photoionization spectra, these authors observed several Rydberg series, including the Kr II $[4s^24p^4(^1D_2)]$ *n*s, *n*d series converging toward the excitation threshold Kr II $[4s^24p^4(^1D_2)]$. Very recently, Sakho, [11] applied the Screening constant by unit nuclear charge (SCUNC) formalism to report precise data belonging to various Rydberg series of Ar II and Kr II as observed in the works of Covington *et al.*, [9] for Ar II and of Hinojosa *et al.*, [10] for Kr II. In the present study, we use the Modified atomic orbital theory [12] [13] [14] [15] [16] to report accurate high lying resonance energies of the Cl I-$[3s^23p^3(^2D_{5/2})]nd$ and $[3s^23p^3(^2P_{3/2})]nd$, Ar II-$3s^23p^4(^1D_2)ns$, *n*d and of the Kr II $[4s^24p^4(^1D_2)]ns$, *n*d and $4s^24p^4(^3P_2,^3P_1)]ns$, $4s^24p^4(^3D_2)]ns$, *n*d and $4s^24p^4(^3D_2,\ ^1S_0)]ns$, *n*d Rydberg series reported. Natural widths of the Ar II-$[3s^23p^4(^1D_2)]$ *n*s, *n*d series are also reported. In Section 2, a brief description of the MAOT formalism is given. The results are present in Section 3. Section 4 concludes the study.

2. Theory

2.1. Brief Description of the Modified Atomic Orbital Theory

In the framework of the modified atomic orbital theory (MAOT), the total energy of a (νl)-given orbital is expressed in the form in Rydberg units

$$E(\nu l) = -\frac{\left[Z - \sigma(l)\right]^2}{\nu^2} \tag{1}$$

In Equation (1), σ is the screening constant relative to the electron occupying the (νl)-orbital, l denotes the orbital quantum number, ν stands for the principal quantum number and Z represents the atomic number. In general, the doubly excited states (DES) in two electron systems are labelled as $(Nl, nl')^{2S+1}L^{\pi}$. In this notation, N and n denote respectively the principal quantum numbers of the inner and the outer electron, l and l' are their respective orbital quantum numbers, S the total spin, L the total angular momentum and π the parity of the system. For an atomic system of many M electrons, total energy is expressed as follows

$$E = -\sum_{i=1}^{M} \frac{\left[Z - \sigma_i\left({}^{2S+1}L^{\pi}\right)\right]^2}{\nu_i^2} \tag{2}$$

In the photoionisation study, energy resonances E_n are generally measured relatively to the E_∞ converging limit of a given $({}^{2S+1}L_J)nl$-Rydberg series. For these states [12] [13] [14]

$$\begin{aligned} E_n = E_\infty - \frac{1}{n^2}\Bigg\{ & Z - \sigma_1\left({}^{2S+1}L_J\right) - \sigma_2\left({}^{2S+1}L_J\right) \times \frac{1}{n} \\ & - \sigma_2^{\alpha}\left({}^2P^0_{3/2}, {}^1D_2\right) \times (n-m) \times (n-q) \sum_k \frac{1}{f_k(n,m,q,s)} \Bigg\}^2 \end{aligned} \tag{3}$$

In this equation, m and q ($m < q$) denote the principal quantum numbers of the $({}^{2S+1}L_J)nl$-Rydberg series of the considered atomic system used in the empirical determination of the $\sigma_i\left({}^{2S+1}L_J\right)$-screening constants, s represents the spin of the nl-electron ($s = 1/2$), E_∞ is the energy value of the series limit generally determined from NIST atomic database, and Z represents the nuclear charge of the considered element. The only problem that one may face by using the MAOT formalism is linked to the determination of the $\sum_k \frac{1}{f_k(n,m,q,s)}$-term. The correct expression of this term is determined iteratively by imposing general Equation (3) to give accurate data with a constant quantum defect values along all the considered series. The value of α is fixed to 1 and or 2 during the iteration. The quantum defect is calculated from the standard formula

$$E_n = E_\infty - \frac{RZ_{\text{core}}^2}{(n-\delta)^2} \tag{4}$$

In this equation, R is the Rydberg constant, E_∞ denotes the converging limit, Z_{core} represents the electric charge of the core ion, and δ means the quantum defect. As far as the natural widths are concerned, they are given by (in Rydberg units)

$$\Gamma_n = \frac{1}{n^2}\left\{Z - \sigma'_1\left({}^{2S+1}L_J\right) - \sigma'_2\left({}^{2S+1}L_J\right)\times\frac{1}{n}\right.$$
$$\left. - \sigma_2^{\alpha'}\left({}^2P^0_{3/2}, {}^1D_2\right)\times(n-m)\times(n-q)\sum_k \frac{1}{f'_k(n,m,q,s)}\right\}^2 \quad (5)$$

2.2. Expressions of the Resonance Energies and of the Natural Widths

In the present work, for all the Rydberg series investigated for both Cl II, Ar II and Kr II, the resonance energies are given by the formula

$$E_n = E_\infty - \frac{1}{n^2}\left\{Z - \sigma_1\left({}^{2S+1}L_J\right) - \sigma_2\left({}^{2S+1}L_J\right)\times\frac{1}{n}\right.$$
$$- \sigma_2^2\left({}^{2S+1}L_J\right)\times(n-m)^2\left[\frac{1}{(n+s+2)^4} + \frac{1}{(n-2s)^4}\right] \quad (6)$$
$$\left. + \sigma_2\left({}^{2S+1}L_J\right)\times(n-m)^2\left[\frac{1}{(n+m+s)^4} + \frac{1}{(n+m-s)^4}\right]\right\}^2$$

For the $[3s^2 3p^4 ({}^1D_2)]ns$ ($j = 1/2$) series originating from the $3s^2 3p^5\ {}^2P^\circ_{1/2}$ metastable state of Ar II ions, the natural widths are given by (in Rydberg units)

$$\Gamma_n = \frac{1}{n^2}\left\{Z - \sigma_1\left({}^{2S+1}L_J\right) - \sigma_2\left({}^{2S+1}L_J\right)\times\frac{1}{n} + \sigma_2^2\left({}^{2S+1}L_J\right)\times(n-m)\times(n-q)\right.$$
$$\left.\times\left[\frac{1}{(n+q-m+s+1)^4} + \frac{1}{(n+m+s)^4} + \frac{1}{(n+q+s)^4} + \frac{1}{(n+m-s)^4}\right]\right\}^2 \quad (7)$$

The other expressions for the other series are of type Equation (7).

3. Results and Discussion

The σ_1-screening constants in Equations (6) and (7) are evaluated empirically using the data from Covington *et al.*, [9] Hinojosa *et al.*, [10] and from Hernández *et al.*, [4]. The results obtained as indicated in the caption of the corresponding Table. As far as the σ_2-screening constant is concerned, it is evaluated theoretical from the simple equation $\sigma_2 = Z - Z_{core}$. The electric charge of the core ion is deduced directly from the single Photoionization process for a given X^{p+}-plasma ion

$$h\nu + X^{p+} \to X^{(p+1)} + e^- \Rightarrow Z_{core} = (p+1) \quad (8)$$

So, for Cl II, Ar II and Kr II, we find respectively.

$$h\nu + Cl^+ \to Cl^{2+} + e^-; Z_{core} = 2 \Rightarrow \sigma_2 = 15.00$$

$$h\nu + Ar^+ \to Ar^{2+} + e^-; Z_{core} = 2 \Rightarrow \sigma_2 = 16.00$$

$$h\nu + Kr^+ \to Kr^{2+} + e^-; Z_{core} = 2 \Rightarrow \sigma_2 = 34.00$$

The resonance energies of the $[3s^2 3p^3\left({}^2D^\circ_{5/2}\right)]nd$ Rydberg series originating from the $3s^2 3p^4\ {}^3P^\circ_2$ ground state and from the $3s^2 3p^4\ {}^3P^\circ_1$, $3s^2 3p^4\ {}^3P^\circ_0$,

$3s^23p^4\,{}^1S_0^\circ$ and $3s^23p^4\,{}^1D_2$ metastable states of Cl II ions are listed in **Tables 1-8**. Comparisons of the present MAOT calculations are done with the available Dirac-Coulomb *R*-matrix (DCR) calculations [5] and with the ALS experimental data [4]. For both the DCR [5] and ALS [4] studies, the determination of the resonances energies have been limited to $n = 13$ (see **Table 4**). In general due to interaction configuration and other electron-electron effects, the peaks of the cross section overlap involving difficulty for the identification of lines in the atomic spectra with increasing n. But, it can be seen that, the MAOT formulas are enough stable so that very high lying resonances can be tabulate up to $n = 40$ with a quantum defect practically constant along all the series investigated. For many resonances, the uncertain experimental entries in parenthesis are enlightened. In **Table 2**, the resonance energy of the $[3s^23p^3\left({}^2D_{5/2}^\circ\right)]11d$ level are equal to 25.493 eV (MAOT), 25.493 eV (DCR) and (25.492 eV) for the ALS uncertain experimental data. The excellent agreement between theories indicate that the ALS data can be stated as accurate at 25.492 eV. The same conclusion can be drawn for the $[3s^23p^3\left({}^2D_{5/2}^\circ\right)]8d$ state quoted in **Table 3** where the MAOT prediction at 25.001 eV agree very well with both the DCR value at 24.999 eV [5] and the uncertain ALS measurement [4] equal to (25.000 eV). For this level the ALS data must be considered as precise at 25.000 eV. **Table 4** lists resonance energies of the $[3s^23p^3\left({}^2P_{3/2}^\circ\right)]nd$ Rydberg series originating from the $3s^23p^4\,{}^3P_2^\circ$ ground state of the Cl^+ ions. In this table, two uncertain ALS values are quoted for the $[3s^23p^3\left({}^2P_{3/2}^\circ\right)]6d$ and $[3s^23p^3\left({}^2P_{3/2}^\circ\right)]12d$ levels respectively at (25.745 eV) and (27.114 eV) to be compared to the MAOT predictions at 25.749 eV and 27.113 eV and to the DCR data [5] respectively equal to 25.755 eV and 27.115 eV. For the $n = 11$ and 13, the MAOT calculations respectively at 27.031 eV and 27.176 eV are seen to agree very well with the ALS measurements [4] at 27.031 eV and 27.175 eV. Subsequently the DCR data at 25.755 eV ($n = 6$) and at 27.178 eV ($n = 13$) are probably greater than the accurate data. The ALS experimental entries in parenthesis can be considered as certain at 25.745 eV and 27.114 eV. Besides, the MAOT data at 25.659 eV quoted in **Table 5** is seen to agree very well with the uncertain ALS measurement [4] at (25.660 eV). A slight discrepancy is observed when comparing with the corresponding DCR calculation [5] equal to 25.668 eV. Comparison indicates clearly that the ALS value [4] is correct at 25.660 eV. In **Table 6** and **Table 7**, all the ALS data [4] are certain. In general, good agreements are obtained between theory and experiment. In **Table 8**, the uncertain ALS data [4] at (24.152 eV) for the $[3s^23p^3\left({}^2D_{5/2}^\circ\right)]11d$ level is difficult to enlighten. For this level, the MAOT prediction at 24.146 eV compared fairly well with the DCR calculations [5] equal to 24.138 eV. A new measurement or calculation is needed to clarify this uncertain ALS value [4]. Overall, for the entire data quoted in **Tables 2-8**, comparisons indicate that the MAOT formula reproduces with a very good accuracy the ALS measurements [4] via a simple formalism without using computational codes in contrast with the DCR formalism [5]. **Tables 9-11** list resonance energies of the

Table 1. Resonance energies of the $[3s^23p^3(^2D^\circ_{5/2})]nd$ Rydberg series originating from the $3s^23p^4\,^3P^\circ_2$ ground state of the Cl$^+$ ions converging to the $3s^23p^3(^2D^\circ_{5/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb *R*-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al*., [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al*., [19]. $\sigma_1(^2D_{5/2}) = -0.770 \pm 0.048$; $\sigma_2(^2D_{5/2}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	E	E	δ	δ	δ
6	24.348	24.353	24.348	0.362	0.30	0.38
7	24.824	24.846	24.829	0.363	0.35	0.35
8	25.127	25.130	25.128	0.362	0.30	0.36
9	25.331	25.334	25.335	0.361	0.34	0.34
10	25.474	25.476	25.479	0.360	0.35	0.32
11	25.579	25.584	25.583	0.360	0.30	0.32
12	25.658			0.359		
13	25.719			0.359		
14	25.768			0.358		
15	25.806			0.358		
16	25.838			0.358		
17	25.863			0.358		
18	25.885			0.358		
19	25.903			0.358		
20	25.919			0.358		
21	25.932			0.359		
22	25.944			0.359		
23	25.954			0.359		
24	25.963			0.359		
25	25.970			0.359		
26	25.977			0.360		
27	25.983			0.360		
28	25.989			0.360		
29	25.994			0.361		
30	25.998			0.361		
31	26.002			0.361		
32	26.006			0.361		
33	26.009			0.362		
34	26.012			0.362		
35	26.015			0.362		
36	26.017			0.363		
37	26.019			0.363		
38	26.022			0.363		
39	26.024			0.363		
40	26.025			0.364		
…	…	…	…		…	…
∞	26.060	26.060	26.060			

Table 2. Resonance energies (E) and quantum defect (δ) of the $[3s^23p^3(^2D^\circ_{5/2})]nd$ Rydberg series originating from the $3s^23p^4\,^3P^\circ_1$ metastable state of the Cl^+ ions converging to the $3s^23p^3(^2D^\circ_{5/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb R-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al.*, [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [19]. $\sigma_1(^2D_{5/2}) = -0.781 \pm 0.048$; $\sigma_2(^2D_{5/2}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	E	E	δ	δ	δ
6	24.259	24.264	24.259	0.367	0.36	0.37
7	24.737	24.762	24.750	0.368	0.30	0.33
8	25.040	25.039	25.036	0.367	0.37	0.38
9	25.244	25.237	25.238	0.366	0.40	0.40
10	25.388	25.385	25.384	0.365	0.40	0.42
11	25.493	25.493	(25.492)	0.365	0.36	(0.37)
12	25.572			0.364		
13	25.633			0.363		
14	25.681			0.363		
15	25.720			0.363		
16	25.751			0.363		
17	25.777			0.363		
18	25.799			0.363		
19	25.817			0.363		
20	25.833			0.363		
21	25.846			0.363		
22	25.858			0.364		
23	25.868			0.364		
24	25.877			0.364		
25	25.884			0.364		
26	25.891			0.365		
27	25.897			0.365		
28	25.903			0.365		
29	25.908			0.365		
30	25.912			0.366		
31	25.916			0.366		
32	25.920			0.366		
33	25.923			0.367		
34	25.926			0.367		
35	25.929			0.367		
36	25.931			0.367		
37	25.933			0.368		
38	25.936			0.368		
39	25.938			0.368		
40	25.939			0.369		
…	…	…	…		…	…
∞	25.974	25.974	25.974			

Table 3. Resonance energies (E) and quantum defect (δ) of the $[3s^2 3p^3(^2D^{\circ}_{5/2})]nd$ Rydberg series originating from the $3s^2 3p^4\,{}^3P^{\circ}_0$ metastable state of the Cl$^+$ ions converging to the $3s^2 3p^3(^2D^{\circ}_{5/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb R-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al.*, [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [19]. $\sigma_1(^2D_{5/2}) = -0.789 \pm 0.048$; $\sigma_2(^2D_{5/2}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	E	E	δ	δ	δ
6	24.219	24.223	24.219	0.370	0.36	0.37
7	24.697	24.733		0.372	0.27	
8	25.001	24.999	(25.000)	0.371	0.38	(0.38)
9	25.205	25.198		0.370	0.40	
10	25.349	25.345		0.369	0.40	
11	25.455			0.368		
12	25.534			0.367		
13	25.595			0.367		
14	25.643			0.367		
15	25.682			0.366		
16	25.713			0.366		
17	25.739			0.366		
18	25.761			0.366		
19	25.779			0.367		
20	25.795			0.367		
21	25.808			0.367		
22	25.820			0.367		
23	25.830			0.367		
24	25.839			0.368		
25	25.846			0.368		
26	25.853			0.368		
27	25.859			0.368		
28	25.865			0.369		
29	25.870			0.369		
30	25.874			0.369		
31	25.878			0.370		
32	25.882			0.370		
33	25.885			0.370		
34	25.888			0.371		
35	25.891			0.371		
36	25.893			0.371		
37	25.895			0.371		
38	25.898			0.372		
39	25.900			0.372		
40	25.901			0.372		
…	…	…	…		…	…
∞	25.936	25.936	25.936			

Table 4. Resonance energies (E) and quantum defect (δ) of the $[3s^2 3p^3(^2P^\circ_{3/2})]nd$ Rydberg series originating from the $3s^2 3p^4\,^3P^\circ_2$ ground state of the Cl^+ ions converging to the $3s^2 3p^3(^2P^\circ_{3/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb R-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al*., [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al*., [19]. $\sigma_1(^2P_{3/2}) = -1.006 \pm 0.077$; $\sigma_2(^2P_{3/2}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	E	E	δ	δ	δ
6	25.749	25.755	(25.745)	0.459	0.46	(0.47)
7	26.246	26.253	26.246	0.469	0.45	0.47
8	26.562	26.576	26.564	0.471	0.45	0.46
9	26.774	26.780	26.778	0.471	0.44	0.45
10	26.923	26.927	26.928	0.470	0.44	0.43
11	27.031	27.035	27.031	0.469	0.43	0.47
12	27.113	27.115	(27.114)	0.468	0.44	(0.45)
13	27.176	27.178	27.175	0.467	0.42	0.48
14	27.225			0.467		
15	27.264			0.466		
16	27.296			0.466		
17	27.323			0.466		
18	27.345			0.466		
19	27.364			0.466		
20	27.379			0.466		
21	27.393			0.466		
22	27.405			0.466		
23	27.415			0.466		
24	27.424			0.467		
25	27.432			0.467		
26	27.439			0.467		
27	27.445			0.467		
28	27.450			0.468		
29	27.455			0.468		
30	27.460			0.468		
31	27.464			0.469		
32	27.467			0.469		
33	27.471			0.470		
34	27.474			0.470		
35	27.476			0.470		
36	27.479			0.471		
37	27.481			0.471		
38	27.483			0.471		
39	27.485			0.472		
40	27.487			0.472		
…	…	…	…		…	…
∞	27.522	27.522	27.522			

Table 5. Resonance energies (E, eV), quantum defect (δ) of the $[3s^23p^3(^2P^\circ_{3/2})]nd$ Rydberg series originating from the $3s^23p^4\,^3P^\circ_1$ metastable state of the Cl^+ ions converging to the $3s^23p^3(^2P^\circ_{3/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb *R*-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al.*, [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [19]. $\sigma_1(^2P_{3/2}) = -1.018 \pm 0.076$; $\sigma_2(^2P_{3/2}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	E	E	δ	δ	δ
6	25.659	25.668	(25.660)	0.464	0.45	(0.45)
7	26.157	26.162	26.157	0.474	0.46	0.47
8	26.473	26.480	26.475	0.477	0.45	0.47
9	26.686	26.695	26.687	0.476	0.43	0.46
10	26.835	26.842	26.833	0.475	0.42	0.49
11	26.944			0.474		
12	27.025			0.473		
13	27.088			0.472		
14	27.138			0.472		
15	27.177			0.471		
16	27.209			0.471		
17	27.236			0.471		
18	27.258			0.471		
19	27.276			0.471		
20	27.292			0.471		
21	27.306			0.471		
22	27.318			0.471		
23	27.328			0.471		
24	27.337			0.472		
25	27.345			0.472		
26	27.351			0.472		
27	27.358			0.473		
28	27.363			0.473		
29	27.368			0.473		
30	27.373			0.474		
31	27.377			0.474		
32	27.380			0.475		
33	27.384			0.475		
34	27.387			0.475		
35	27.389			0.476		
36	27.392			0.476		
37	27.394			0.476		
38	27.396			0.477		
39	27.398			0.477		
40	27.400			0.477		
…	…	…	…		…	…
∞	27.435	27.435	27.435			

Table 6. Resonance energies (*E*) and quantum defect (δ) of the $[3s^2 3p^3(^2P^{\circ}_{3/2})]nd$ Rydberg series originating from the $3s^2 3p^4\,^3P^{\circ}_0$ metastable state of the Cl^+ ions converging to the $3s^2 3p^3(^2P^{\circ}_{3/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb *R*-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al*., [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al*., [19]. $\sigma_1(^2P_{3/2}) = -1.088 \pm 0.076$; $\sigma_2(^2P_{32}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	*E*	*E*	δ	δ	δ
6	25.603	25.635		0.493	0.44	
7	26.108	26.117	26.108	0.505	0.48	0.51
8	26.429	26.434	26.430	0.507	0.49	0.50
9	26.644	26.650	26.648	0.507	0.47	0.49
10	26.794	26.797		0.506	0.48	
11	26.904			0.504		
12	26.986			0.503		
13	27.050			0.502		
14	27.099			0.502		
15	27.139			0.501		
16	27.171			0.501		
17	27.198			0.501		
18	27.220			0.501		
19	27.239			0.501		
20	27.255			0.501		
21	27.268			0.501		
22	27.280			0.501		
23	27.290			0.502		
24	27.299			0.502		
25	27.307			0.502		
26	27.314			0.503		
27	27.320			0.503		
28	27.326			0.504		
29	27.331			0.504		
30	27.335			0.504		
31	27.339			0.505		
32	27.343			0.505		
33	27.346			0.506		
34	27.349			0.506		
35	27.352			0.507		
36	27.355			0.507		
37	27.357			0.507		
38	27.359			0.508		
39	27.361			0.508		
40	27.363			0.509		
…	…	…	…		…	…
∞	27.398	27.398	27.398			

Table 7. Resonance energies (E, eV), quantum defect (δ) of the $[3s^2 3p^3(^2P^\circ_{1/2})]nd$ Rydberg series originating from the $3s^2 3p^4\,^1S^\circ_0$ metastable state of the Cl^+ ions converging to the $3s^2 3p^3(^2P^\circ_{1/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb *R*-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al.*, [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [19]. $\sigma_1(^2P_{1/2}) = -0.262 \pm 0.015$; $\sigma_2(^2P_{1/2}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	E	E	δ	δ	δ
4	20.426	20.420	20.426	0.127	0.13	0.13
5	21.762	21.734	21.742	0.127	0.16	0.15
6	22.477	22.459	22.463	0.126	0.16	0.15
7	22.902	22.890	22.900	0.126	0.16	0.13
8	23.176	23.170	23.178	0.125	0.15	0.12
9	23.363	23.361		0.125	0.14	
10	23.496			0.125		
11	23.594			0.124		
12	23.668			0.124		
13	23.726			0.124		
14	23.771			0.124		
15	23.808			0.124		
16	23.838			0.124		
17	23.863			0.124		
18	23.884			0.124		
19	23.901			0.124		
20	23.916			0.124		
21	23.929			0.124		
22	23.940			0.124		
23	23.950			0.124		
24	23.959			0.124		
25	23.966			0.124		
26	23.973			0.125		
27	23.979			0.125		
28	23.984			0.125		
29	23.989			0.125		
30	23.993			0.125		
31	23.997			0.125		
32	24.000			0.125		
33	24.004			0.125		
34	24.007			0.125		
35	24.009			0.125		
36	24.012			0.125		
37	24.014			0.125		
…	…	…	…		…	…
∞	24.054	24.054	24.054			

Table 8. Resonance energies (E) and quantum defect (δ) of the $[3s^2 3p^3(^2D^{\circ}_{5/2})]nd$ series originating from the $3s^2 3p^4\,{}^1D_2$ metastable state of the Cl$^+$ ions converging to the $3s^2 3p^3(^2D^{\circ}_{5/2})$ threshold of Cl^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Dirac-Coulomb R-matrix (DCR) calculations of McLaughlin [5] and with the ALS experimental data of Hernández *et al.*, [4]. The ALS experimental resonance energies are calibrated to ±0.013 eV. The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [19]. $\sigma_1(^2D_{5/2}) = -0.483 \pm 0.050$; $\sigma_2(^2D_{5/2}) = 15.00$.

n	MAOT	DCR	ALS	MAOT	DCR	ALS
	E	E	E	δ	δ	δ
6	22.979	22.969	22.979	0.232	0.25	0.27
7	23.427	23.392		0.233	0.32	
8	23.713	23.719	23.718	0.232	0.21	0.21
9	23.907	23.907	23.907	0.232	0.23	0.23
10	24.045	24.039	24.040	0.231	0.28	0.27
11	24.146	24.138	(24.152)	0.231	0.32	(0.27)
12	24.222			0.230		
13	24.281			0.230		
14	24.328			0.230		
15	24.366			0.230		
16	24.396			0.230		
17	24.421			0.229		
18	24.443			0.229		
19	24.461			0.229		
20	24.476			0.229		
21	24.489			0.229		
22	24.500			0.229		
23	24.510			0.229		
24	24.519			0.230		
25	24.526			0.230		
26	24.533			0.230		
27	24.539			0.230		
28	24.544			0.230		
29	24.549			0.230		
30	24.554			0.230		
31	24.558			0.230		
32	24.561			0.230		
33	24.564			0.230		
34	24.567			0.231		
35	24.570			0.231		
36	24.572			0.231		
37	24.575			0.231		
38	24.577			0.231		
39	24.579			0.231		
40	24.581			0.231		
…	…	…	…		…	…
∞	24.615	24.615	24.615			

Table 9. Resonance energies (E) and quantum defect (δ) of the $3s^23p^4(^1D_2)ns$ ($j = 1/2$ series originating from the $3s^23p^5\,^2P^{\circ}_{1/2}$ metastable state of the Ar^+ ions converging to the $^1D_2$ threshold of Ar^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al*., [9]. The energy limits eV) is taken from the NIST tabulations of Ralchenko *et al*., [17]. $\sigma_1(^1D_2) = -4.227 \pm 0.224$; $\sigma_2(^1D_2) = 16.00$.

n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	E	E	E	δ	δ	δ
8	27.830	27.830	27.830	1.672	1.673	1.673
9	28.169	28.173	28.173	1.695	1.682	1.682
10	28.400	28.401	28.401	1.696	1.689	1.688
11	28.561	28.561	28.561	1.689	1.694	1.692
12	28.678	28.676	28.677	1.679	1.697	1.695
13	28.765	28.763	28.763	1.670	1.699	1.698
14	28.831	28.829	28.829	1.662	1.700	1.699
15	28.883	28.881	28.881	1.656	1.701	1.701
16	28.925	28.923	28.923	1.652	1.701	1.702
17	28.958	28.956		1.649	1.700	
18	28.985	28.984		1.647	1.700	
19	29.008	29.007		1.647	1.698	
20	29.027	29.027		1.648	1.697	
21	29.044	29.043		1.649	1.695	
22	29.058	29.057		1.652	1.693	
23	29.070	29.069		1.655	1.691	
24	29.080	29.080		1.658	1.688	
25	29.089	29.089		1.662	1.686	
26	29.097	29.097		1.666	1.683	
27	29.104	29.104		1.670	1.680	
28	29.110	29.110		1.675	1.677	
29	29.116	29.116		1.679	1.674	
30	29.121	29.121		1.684	1.671	
31	29.126			1.689		
32	29.130			1.694		
33	29.133			1.699		
34	29.137			1.704		
35	29.140			1.708		
36	29.143			1.713		
37	29.145			1.718		
38	29.148			1.723		
39	29.150			1.727		
40	29.152			1.732		
…	…	…	…		…	…
∞	29.189	29.189	29.189			

Table 10. Resonance energies (E) and quantum defect (δ) of the $3s^23p^4$ ($^1D_2$)nd ($j = 1/2$ series originating from the $3s^23p^{5\ 2}P^{\circ}_{1/2}$ metastable state of the Ar^+ ions converging to the $3s^23p^4(^1D_2)$ threshold of Ar^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al.*, [9]. The energy limits eV) is taken from the NIST tabulations of Ralchenko *et al.*, [17]. $\sigma_1(^1D_2) = -1.159 \pm 0.265$; $\sigma_2(^1D_2) = 16.00$.

n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	E	E	E	δ	δ	δ
8	28.211	28.211	28.211	0.540	0.540	0.540
9	28.428	28.426	28.426	0.543	0.555	0.554
10	28.581	28.578	28.576	0.543	0.563	0.574
11	28.691	28.689	28.691	0.542	0.568	0.546
12	28.775	28.772	28.773	0.541	0.571	0.562
13	28.838	28.837	28.837	0.540	0.573	0.568
14	28.889	28.887	28.887	0.539	0.573	0.572
15	28.929	28.928	28.928	0.538	0.573	0.574
16	28.961	28.960		0.538	0.572	
17	28.988	28.987		0.537	0.571	
18	29.011	29.010		0.537	0.570	
19	29.029	29.029		0.537	0.568	
20	29.045	29.045		0.537	0.567	
21	29.059	29.059		0.537	0.565	
22	29.071	29.071		0.537	0.563	
23	29.081	29.081		0.537	0.560	
24	29.090	29.090		0.537	0.558	
25	29.098	29.098		0.538	0.556	
26	29.105	29.105		0.538	0.553	
27	29.111	29.111		0.538	0.551	
28	29.117	29.117		0.538	0.548	
29	29.122	29.122		0.539	0.546	
30	29.126	29.126		0.539	0.543	
31	29.130			0.539		
32	29.134			0.540		
33	29.137			0.540		
34	29.140			0.540		
35	29.143			0.541		
36	29.146			0.541		
37	29.148			0.542		
38	29.150			0.542		
39	29.152			0.542		
40	29.154			0.543		
…	…	…	…		…	…
∞	29.189	29.189	29.189			

Table 11. Resonance energies (E) and quantum defect (δ) of the $3s^23p^4(^1D_2)nd$ ($j = 3/2$ series originating from the $3s^23p^5\,^2P^{\circ}_{1/2}$ metastable state of the Ar^+ ions converging to the $3s^23p^4(^1D_2)$ threshold of Ar^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al.*, [9]. The energy limits eV) is taken from the NIST tabulations of Ralchenko *et al.*, [17]. $\sigma_1(^1D_2) = -1.159 \pm 0.265$; $\sigma_2(^1D_2) = 16.00$.

n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	E	E	E	δ	δ	δ
8	27.821	27.821	27.821	1.693	1.693	1.692
9	28.163	28.171	28.171	1.717	1.688	1.691
10	28.396	28.401	28.401	1.718	1.689	1.690
11	28.558	28.561	28.561	1.710	1.692	1.689
12	28.676	28.677	28.677	1.700	1.693	1.688
13	28.763	28.763	28.764	1.691	1.695	1.688
14	28.830	28.830	28.830	1.683	1.695	1.688
15	28.882	28.882	28.882	1.676	1.695	1.687
16	28.924	28.923	28.923	1.672	1.694	1.687
17	28.957	28.957		1.669	1.692	
18	28.985	28.984		1.667	1.691	
19	29.008	29.007		1.667	1.688	
20	29.027	29.027		1.668	1.686	
21	29.043	29.043		1.669	1.683	
22	29.057	29.057		1.672	1.679	
23	29.069	29.069		1.675	1.676	
24	29.080	29.080		1.678	1.672	
25	29.089	29.089		1.682	1.668	
26	29.097	29.097		1.686	1.663	
27	29.104	29.104		1.691	1.659	
28	29.110	29.111		1.695	1.655	
29	29.116	29.116		1.700	1.650	
30	29.121	29.121		1.705	1.645	
31	29.126			1.710		
32	29.130			1.715		
33	29.133			1.720		
34	29.137			1.725		
35	29.140			1.730		
36	29.143			1.735		
37	29.145			1.740		
38	29.148			1.745		
39	29.150			1.750		
40	29.152			1.754		
…	…	…	…		…	…
∞	29.189	29.189	29.189			

$3s^23p^4(^1D_2)ns$, nd (j = 1/2) and of the $3s^23p^4(^1D_2)nd$ (j = 3/2) series originating from the $3s^23p^5\,^2P^\circ_{1/2}$ metastable state of the Ar II ions. Comparison indicate an excellent agreements between the present results from the Modified atomic orbital theory (MAOT) and both the Screening constant by unit nuclear charge (SCUNC) results [11] and the multichannel *R*-matrix QB technique which defines matrices *Q* and *B* in terms of asymptotic solutions [9]. These very good agreements are also observed comparing the resonance energies $3s^23p^4(^1D_2)ns$, nd (j = 1/2) series originating from the $3s^23p^5\,^2P^\circ_{3/2}$ ground state of the Ar II ions. For both **Tables 9-12**, the QB data are limited to n = 16 and the SCUNC values to n = 30. High lying MAOT data are tabulated up to n = 40 with a constant quantum defect along each series. **Table 13** and **Table 14** quote respectively resonance energies of the $3s^23p^4(^1D_2)nd$ (j = 1/2)and of the $3s^23p^4(^1D_2)ns$ (j = 3/2) series originating from the $3s^23p^5\,^2P^\circ_{3/2}$ ground state of the Ar II ions. Here again, the agreements between the present MAOT results and both SCUNC of Sakho [11] and QB results of Covington *et al*., [9] are very good. **Table 15** presents resonance energies of the $3s^23p^4$ $(^1D_2)nd$ (j = 3/2) series originating from the $3s^23p^5\,^2P^\circ_{3/2}$ ground state of the Ar II ions. In this Table, an uncertain QB data [9] is quoted at (28.774 eV) for the $3s^23p^4(^1D_2)10d$ level. For this level, the MOAT prediction is at 28.735 eV to be compared to the SCUNC forecast at 28.734 eV. Subsequently, the QB data at (28.774 eV) associated with a quantum defect equal to (0.422) is less precise. For this level, both the MAOT and SCUNC [11] predictions associated with the quantum defects 0.721 and 0.724 are preferable. Therefore, it should be underlined that the SCUNC calculations are more precise than the MOAT calculations. This is due mainly to the fact that, the SCUNC formalism is a development of $1/Z$ taking implicitly into account more relativistic effects than the MAOT formalism witch is a simple development on $1/n$. In addition, in the SCUNC formalism, great accuracy are obtained when performing the analytical formula for each atomic system [11]. In the present work, the same Formula (6) is used for both Cl II, Ar II and Kr II in contrast with the work of Sakho [11] where the resonance energy expression for the Ar II ions is different to that of the Kr II ions. **Table 16** and **Table 17** list natural widths of the $[3s^23p^4(^1D_2)]ns$, nd (j = 1/2) (**Table 16**) and of the $[3s^23p^4(^1D_2)]ns$, nd (j = 3/2) (**Table 17**) series originating from the $3s^23p^5\,^2P^\circ_{1/2}$ metastable state of Ar^+ ions. It can be seen that the present MAOT data agree well with both the SCUNC results [11] and QB data [9]. It should be mentioned again that the SCUNC calculations are more precise than the MOAT calculations for the reason explained above. **Table 18** lists resonance energies and quantum defect of the $[4s^24p^4$ $(^1D_2)]nd$ series originating from the $4s^24p^5\,^2P^\circ_{3/2}$ ground state of the Kr II ions. Comparisons indicate very good agreements between the present results from the MAOT formalism and both the SCUNC calculations [11] up to n = 30 and ALS measurements of [10] up to n = 13. In this iable, two uncertain ALS data are quoted for the $[4s^24p^4(^1D_2)]12d$ and $[4s^24p^4(^1D_2)]13d$ levels respectively at (25.880 eV) and (25.926 eV). The

Table 12. Resonance energies (E) and quantum defect (δ) of the $3s^23p^4(^1D_2)ns$ ($j = 1/2$ series originating from the $3s^23p^5\,^2P^\circ_{3/2}$ ground state of the Ar^+ ions converging to the $3s^23p^4(^1D_2)$ threshold of Ar^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al.*, [9]. The energy limits eV) is taken from the NIST tabulations of Ralchenko *et al.*, [17]. $\sigma_1(^1D_2) = -4.234 \pm 0.224$; $\sigma_2(^1D_2) = 16.00$.

n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	E	E	E	δ	δ	δ
8	28.007	28.007	28.007	1.674	1.674	1.673
9	28.346	28.351	28.351	1.698	1.681	1.682
10	28.577	28.579	28.579	1.699	1.688	1.688
11	28.739	28.739	28.739	1.691	1.694	1.692
12	28.856	28.854	28.854	1.681	1.698	1.696
13	28.943	28.941	28.941	1.672	1.701	1.698
14	29.009	29.007	29.007	1.664	1.702	1.700
15	29.061	29.059	29.059	1.658	1.702	1.701
16	29.103	29.101	29.100	1.654	1.702	1.702
17	29.136	29.135		1.651	1.700	
18	29.163	29.162		1.649	1.698	
19	29.186	29.185		1.649	1.696	
20	29.205	29.205		1.650	1.693	
21	29.222	29.221		1.652	1.690	
22	29.236	29.235		1.654	1.686	
23	29.248	29.247		1.657	1.682	
24	29.258	29.258		1.660	1.678	
25	29.267	29.267		1.664	1.673	
26	29.275	29.275		1.668	1.668	
27	29.282	29.282		1.672	1.663	
28	29.288	29.289		1.677	1.658	
29	29.294	29.294		1.682	1.653	
30	29.299	29.299		1.686	1.647	
31	29.304			1.691		
32	29.308			1.696		
33	29.311			1.701		
34	29.315			1.706		
35	29.318			1.711		
36	29.321			1.716		
37	29.323			1.720		
38	29.326			1.725		
39	29.328			1.730		
40	29.330			1.734		
…	…	…	…		…	…
∞	29.367	29.367	29.367			

Table 13. Resonance energies (*E*) and quantum defect (*δ*) of the $3s^23p^4(^1D_2)nd$ ($j = 1/2$ series originating from the $3s^23p^5\,^2P^{\circ}_{3/2}$ ground state of the Ar^+ ions converging to the $3s^23p^4(^1D_2)$ threshold of Ar^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al*., [9]. The energy limits eV) is taken from the NIST tabulations of Ralchenko *et al*., [17]. $\sigma_1(^1D_2) = -1.159 \pm 0.265$; $\sigma_2(^1D_2) = 16.00$.

n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	E	*E*	*E*	*δ*	*δ*	*δ*
8	28.389	28.389	28.389	0.540	0.540	0.540
9	28.606	28.603	28.603	0.543	0.560	0.554
10	28.759	28.755	28.754	0.543	0.567	0.574
11	28.869	28.867	28.869	0.542	0.572	0.546
12	28.953	28.950	28.951	0.541	0.575	0.562
13	29.016	29.014	29.014	0.540	0.578	0.568
14	29.067	29.065	29.065	0.539	0.581	0.572
15	29.107	29.105	29.105	0.538	0.583	0.574
16	29.139	29.138		0.538	0.585	
17	29.166	29.165		0.537	0.587	
18	29.189	29.187		0.537	0.588	
19	29.207	29.206		0.537	0.589	
20	29.223	29.223		0.537	0.590	
21	29.237	29.236		0.537	0.589	
22	29.249	29.248		0.537	0.589	
23	29.259	29.259		0.537	0.587	
24	29.268	29.268		0.537	0.585	
25	29.276	29.276		0.538	0.581	
26	29.283	29.283		0.538	0.577	
27	29.289	29.289		0.538	0.572	
28	29.295	29.295		0.538	0.566	
29	29.300	29.300		0.539	0.559	
30	29.304	29.304		0.539	0.551	
31	29.308			0.539		
32	29.312			0.540		
33	29.315			0.540		
34	29.318			0.540		
35	29.321			0.541		
36	29.324			0.541		
37	29.326			0.542		
38	29.328			0.542		
39	29.330			0.542		
40	29.332			0.543		
…	…	…	…		…	…
∞	29.367	29.367	29.367			

Table 14. Resonance energies (*E*) and quantum defect (δ) of the $3s^2 3p^4(^1D_2)ns$ ($j = 3/2$ series originating from the $3s^2 3p^5\,^2P^{\circ}_{3/2}$ ground state of the Ar^+ ions converging to the $3s^2 3p^4(^1D_2)$ threshold of Ar^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al.*, [9]. The energy limits eV) is taken from the NIST tabulations of Ralchenko *et al.*, [17]. $\sigma_1(^1D_2) = -4.294 \pm 0.224$; $\sigma_2(^1D_2) = 16.00$.

n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	E	*E*	*E*	δ	δ	δ
8	27.999	27.999	27.999	1.693	1.693	1.692
9	28.341	28.348	28.348	1.717	1.692	1.691
10	28.574	28.578	28.579	1.718	1.692	1.690
11	28.736	28.739	28.739	1.710	1.693	1.689
12	28.854	28.855	28.855	1.700	1.694	1.688
13	28.941	28.941	28.941	1.691	1.694	1.688
14	29.008	29.008	29.098	1.683	1.695	1.688
15	29.060	29.060	29.060	1.676	1.695	1.687
16	29.102	29.101	29.101	1.672	1.695	1.687
17	29.135	29.135		1.669	1.695	
18	29.163	29.162		1.667	1.695	
19	29.186	29.185		1.667	1.694	
20	29.205	29.205		1.668	1.694	
21	29.221	29.221		1.669	1.693	
22	29.235	29.235		1.672	1.692	
23	29.247	29.247		1.675	1.691	
24	29.258	29.258		1.678	1.690	
25	29.267	29.267		1.682	1.689	
26	29.275	29.275		1.686	1.688	
27	29.282	29.282		1.691	1.687	
28	29.288	29.288		1.695	1.686	
29	29.294	29.294		1.700	1.684	
30	29.299	29.299		1.705	1.683	
31	29.304			1.710		
32	29.308			1.715		
33	29.311			1.720		
34	29.315			1.725		
35	29.318			1.730		
36	29.321			1.735		
37	29.323			1.740		
38	29.326			1.745		
39	29.328			1.750		
40	29.330			1.754		
…	…	…	…		…	…
∞	29.367	29.367	29.367			

Table 15. Resonance energies (E, eV) and quantum defect (δ) of the $3s^23p^4(^1D_2)nd$ (j = 3/2 series originating from the $3s^23p^5\,^2P^{\circ}_{3/2}$ ground state of the Ar^+ ions converging to the $3s^23p^4(^1D_2)$ threshold of Ar^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al.*, [9]. The energy limits eV) is taken from the NIST tabulations of Ralchenko *et al.*, [17]. $\sigma_1(^1D_2) = -1.575 \pm 0.265$; $\sigma_2(^1D_2) = 16.00$.

n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	E	E	E	δ	δ	δ
8	28.341	28.341	28.341	0.717	0.717	0.716
9	28.573	28.573	28.573	0.721	0.721	0.720
10	28.735	28.734	(28.774)	0.721	0.724	(0.422)*
11	28.852	28.851	28.850	0.720	0.727	0.738
12	28.939	28.939	28.937	0.718	0.729	0.745
13	29.006	29.006	29.004	0.717	0.729	0.744
14	29.059	29.058	29.047	0.715	0.729	0.744
15	29.100	29.100	29.099	0.714	0.728	0.744
16	29.134	29.134		0.713	0.727	
17	29.162	29.162		0.713	0.725	
18	29.185	29.185		0.713	0.722	
19	29.204	29.204		0.712	0.720	
20	29.221	29.221		0.712	0.718	
21	29.235	29.235		0.713	0.716	
22	29.247	29.247		0.713	0.714	
23	29.257	29.257		0.713	0.713	
24	29.267	29.267		0.714	0.712	
25	29.275	29.275		0.714	0.712	
26	29.282	29.282		0.715	0.713	
27	29.288	29.288		0.715	0.715	
28	29.294	29.294		0.716	0.719	
29	29.299	29.299		0.717	0.723	
30	29.304	29.303		0.717	0.729	
31	29.308			0.718		
32	29.311			0.719		
33	29.315			0.719		
34	29.318			0.720		
35	29.321			0.721		
36	29.323			0.722		
37	29.326			0.722		
38	29.328			0.723		
39	29.330			0.724		
40	29.332			0.724		
…	…	…	…		…	…
∞	29.367	29.367	29.367			

*This line is not well identified.

Table 16. Natural widths (Γ, meV) of the $[3s^23p^4(^1D_2)]ns$, nd ($j = 1/2$) series originating from the $3s^23p^5\,^2P^\circ_{1/2}$ metastable state of Ar$^+$ ions. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al.*, [9]. For: $[4s^24p^4(^1D_2)]$ ns: $\sigma_1(^1D_2) = -5.494 \pm 0.010$; $\sigma_2(^1D_2) = 17.989 \pm 0.010$. For: $[4s^24p^4(^1D_2)]nd$: $\sigma_1(^1D_2) = -0.168 \pm 0.010$; $\sigma_2(^1D_2) = 17.998 \pm 0.010$.

	$[4s^24p^4(^1D_2)]ns$ series			$[4s^24p^4(^1D_2)]nd$ series		
n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	Γ	Γ	Γ	Γ	Γ	Γ
8	103.5	103.4	103.4	0.112	0.110	0.110
9	64.9	64.8	64.8	0.072	0.070	0.070
10	43.4	43.0	43.0	0.048	0.047	0.050
11	30.4	30.0	30.0	0.034	0.033	0.040
12	22.1	21.7	21.8	0.024	0.024	0.030
13	16.5	16.2	16.3	0.018	0.018	0.020
14	12.7	12.4	12.5	0.014	0.014	0.020
15	9.9	9.8	9.8	0.011	0.011	0.020
16	7.8	7.8	7.8	0.008	0.009	
17	6.3	6.4		0.007	0.008	
18	5.1	5.3		0.005	0.006	
19	4.2	4.5		0.004	0.005	
20	3.5	3.8		0.004	0.005	
21	2.9	3.3		0.003	0.004	
22	2.5	2.9		0.003	0.003	
23	2.1	2.5		0.002	0.003	
24	1.8	2.2		0.002	0.003	
25	1.5	2.0		0.002	0.002	

Table 17. Natural widths (Γ, meV) of the $[3s^23p^4(^1D_2)]ns$, nd ($j = 3/2$) series originating from the $3s^23p^5\,^2P^\circ_{1/2}$ metastable state of Ar$^+$ ions. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the QB data of Covington *et al.*, [9]. For: $[4s^24p^4(^1D_2)]$ ns: $\sigma_1(^1D_2) = -0.670 \pm 0.010$; $\sigma_2(^1D_2) = 18.038 \pm 0.010$. For: $[4s^24p^4(^1D_2)]$ nd: $\sigma_1(^1D_2) = -1.399 \pm 0.010$; $\sigma_2(^1D_2) = 17.827 \pm 0.010$.

	$[4s^24p^4(^1D_2)]ns$ series			$[4s^24p^4(^1D_2)]nd$ series		
n	MAOT	SCUNC	QB	MAOT	SCUNC	QB
	Γ	Γ	Γ	Γ	Γ	Γ
8	0.44	0.44	0.44	25.7		25.7
9	0.22	0.22	0.22	18.1		18.1
10	0.13	0.11	0.13	12.9		33.8
11	0.08	0.06	0.08	9.3		8.5

Continued

12	0.05	0.03	0.08	6.9		3.8
13	0.03	0.02	0.04	5.2		5.3
14	0.02	0.01	0.02	4.0		4.4
15	0.01	0.01	0.02	3.1		3.6
16				2.5		
17				2.1		
18				1.7		
19				1.4		
20				1.2		
21				1.0		
22				0.9		
23				0.8		
24				0.7		
25				0.6		

Table 18. Resonance energies (E) and quantum defect (δ) of the $[4s^2 4p^4(^1D_2)]nd$ series originating from the $4s^2 4p^5\,{}^2P^{\circ}_{3/2}$ ground state of the Kr^+ ions converging to the $[4s^2 4p^4(^1D_2)]$ threshold of Kr^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the ALS experimental data of Hinojoha *et al.*, [10]. The ALS resonance energies are calibrated to ±30 meV and quantum defects are estimated to within an error of 20%. The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [18]. Here $\sigma_1(^1D_2) = -0.719 \pm 0.116$; $\sigma_2(^1D_2) = 34.00$.

n	MAOT	SCUNC	ALS	MAOT	SCUNC	ALS
	E	E	E	δ	δ	δ
5	24.342	24.342	24.342	−0.387	−0.387	−0.385
6	24.882	24.878	24.878	−0.386	−0.375	−0.370
7	25.220	25.215	25.217	−0.387	−0.369	−0.370
8	25.444	25.439	25.441	−0.389	−0.366	−0.370
9	25.600	25.596	25.598	−0.390	−0.365	−0.360
10	25.713	25.710	25.712	−0.390	−0.365	−0.358
11	25.798	25.796	25.796	−0.391	−0.365	−0.345
12	25.863	25.861	(25.880)	−0.391	−0.366	(−0.671)
13	25.913	25.912	(25.926)	−0.391	−0.367	(−0.663)
14	25.954	25.953		−0.391	−0.368	
15	25.987	25.987		−0.390	−0.369	
16	26.014	26.014		−0.390	−0.370	
17	26.037	26.037		−0.390	−0.371	
18	26.056	26.056		−0.389	−0.372	
19	26.072	26.072		−0.389	−0.373	
20	26.086	26.086		−0.389	−0.373	
21	26.098	26.098		−0.388	−0.374	
22	26.108	26.108		−0.388	−0.374	
23	26.118	26.117		−0.387	−0.374	
24	26.125	26.125		−0.387	−0.374	

Continued

25	26.133	26.132		−0.387	−0.374	
26	26.139	26.139		−0.386	−0.374	
27	26.144	26.144		−0.386	−0.374	
28	26.149	26.149		−0.385	−0.373	
29	26.154	26.154		−0.385	−0.373	
30	26.158	26.158		−0.385	−0.372	
31	26.162			−0.384		
32	26.165			−0.384		
33	26.168			−0.383		
34	26.171			−0.383		
35	26.174			−0.383		
36	26.176			−0.382		
37	26.178			−0.382		
38	26.180			−0.382		
39	26.182			−0.381		
40	26.184			−0.381		
…	…	…	…		…	…
∞	29.367	29.367	29.367			

associated ALS quantum defects are equal to (−0.671) and (−0.663) respectively. For the same levels, the MAOT and SCUNC calculations [11] are respectively at 25.863 eV and 25.861 eV for the $[4s^24p^4(^1D_2)]12d$ state and at 25.913 eV and 25.912 eV for the $[4s^24p^4(^1D_2)]13d$ state. Constant quantum defects are tabulated for the MAOT and SCUNC predictions [11] respectively at −0.31 and −0.37. So the MAOT and SCUNC estimations can be considered as the accurate data for the $[4s^24p^4(^1D_2)]12d$ and $[4s^24p^4(^1D_2)]13d$ resonances. **Table 19** quotes resonance energies and quantum defect of the $4s^24p^4(^3P_2,{}^3P_1)]ns$ and $4s^24p^4(^3P_1)]ns$ series originating from the $4s^24p^5\,{}^2P^{\circ}_{1/2}$ metastable state of the Kr II ions. For the $4s^24p^4(^3P_2)]ns$ series, comparisons indicate good agreements between theory and experiments. It should be underlined the very good agreements for n = 13 - 20. This may enlighten the accuracy of the uncertain ALS measurement [10] listed into parenthesis. For the $4s^24p^4(^3P_1)]ns$ series only one ALS data at 23.996 eV is quoted. New MAOT values from n = 14 to 40 are tabulated with a constant quantum defect about 0.42. In **Table 20**, resonance energies and quantum defect of the $4s^24p^4(^3D_2)]ns$, $n$dseries originating from the $4s^24p^5\,{}^2P^{\circ}_{1/2}$ metastable state of the Kr II ions are listed. Here again, the MAOT data agree very well with the ALS data of Hinojoha *et al.*, [10]. It should be underlined the excellent agreements between theory and experiments for the $4s^24p^4(^3D_2)]8s$ and $4s^24p^4(^3D_2)]9s$ levels. For these states, both the MAOT and ALS work provide the same values respectively equal to 24.650 eV and 24.842 eV. The excellent agreement for the quantum defects can also be mentioned, 0.20 for both theory and experiments. **Table 21** presents resonance energies and quantum defect for the

Table 19. Resonance energies (E) and quantum defect (δ) of the $4s^24p^4(^3P_2,^3P_1)]ns$ series originating from the $4s^24p^5\,^2P^\circ_{1/2}$ metastable state of the Kr$^+$ ions converging to the $[4s^24p^4(^1D_2)]$ threshold of Kr^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the ALS experimental data of Hinojoha *et al.*, [10]. The ALS resonance energies are calibrated to ±30 meV and quantum defects are estimated to within an error of 20%.The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [18]. $\sigma_1(^3P_2) = -1.438 \pm 0.104$; $\sigma_2(^3P_2) = 34.00$; $\sigma_1(^3P_1) = -0.863 \pm 0.138$; $\sigma_2(^3P_1) = 34.00$.

n	$4s^24p^4(^3P_2)]ns$				$4s^24p^4(^3P_1)]ns$			
	MAOT	ALS	MAOT	ALS	MAOT	ALS	MAOT	ALS
	E	E	δ	δ	E	E	δ	δ
11	23.910	(23.906)	−0.771	(−0.716)				
12	23.969	23.969	−0.765	−0.779				
13	24.016	24.016	−0.762	−0.791				
14	24.053	(24.059)	−0.761	(−0.940)	23.957		0.418	
15	24.084	24.083	−0.761	−0.730	23.996	23.996	0.419	0.403
16	24.109	24.109	−0.762	−0.774	24.028		0.420	
17	24.131	(24.129)	−0.762	(−0.686)	24.054		0.420	
18	24.148	(24.147)	−0.763	(−0.707)	24.076		0.420	
19	24.164	(24.162)	−0.764	(−0.659)	24.094		0.420	
20	24.177	(24.176)	−0.764	(−0.731)	24.110		0.420	
21	24.188		−0.765		24.124		0.420	
22	24.198		−0.765		24.135		0.420	
23	24.207		−0.765		24.145		0.419	
24	24.214		−0.766		24.154		0.419	
25	24.221		−0.766		24.162		0.419	
26	24.227		−0.766		24.169		0.419	
27	24.232		−0.766		24.175		0.419	
28	24.237		−0.766		24.180		0.419	
29	24.242		−0.766		24.185		0.418	
30	24.246		−0.766		24.190		0.418	
31	24.249		−0.766		24.194		0.418	
32	24.252		−0.765		24.197		0.418	
33	24.255		−0.765		24.201		0.418	
34	24.258		−0.765		24.204		0.418	
35	24.260		−0.765		24.206		0.418	
36	24.263		−0.765		24.209		0.418	
37	24.265		−0.764		24.211		0.418	
38	24.267		−0.764		24.213		0.418	
39	24.269		−0.764		24.215		0.418	
40	24.270		−0.763		24.217		0.418	
…								
∞	24.303	24.303			24.252	24.252		

Table 20. Resonance energies (*E*) and quantum defect (*δ*) of the $4s^24p^4(^3D_2)]ns$, *n*d series originating from the $4s^24p^5\,{}^2P^\circ_{1/2}$ metastable state of the Kr^+ ions converging to the $[4s^24p^4(^1D_2)]$ threshold of Kr^{2+}. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the ALS experimental data of Hinojosa *et al.*, [10]. The ALS resonance energies are calibrated to ±30 meV and quantum defects are estimated to within an error of 20%. The energy limits is taken from the NIST tabulations of Ralchenko *et al.*, [18]. $\sigma_1(^1D_2) = -0.414 \pm 0.116$; $\sigma_2(^1D_2) = 34.00$ for $4s^24p^4(^3D_2)]ns$. $\sigma_1(^1D_2) = -0.698 \pm 0.116$; $\sigma_2(^1D_2) = 34.00$ for $4s^24p^4(^3D_2)]nd$.

n	$4s^24p^4(^3D_2)]ns$				$4s^24p^4(^3D_2)]nd$			
	MAOT	ALS	MAOT	ALS	MAOT	ALS	MAOT	ALS
	E	*E*	*δ*	*δ*	*E*	*E*	*δ*	*δ*
6	23.927	23.927	0.200	0.200	24.214	24.214	−0.371	−0.375
7	24.368	(24.366)	0.200	(0.205)	24.553	24.551	−0.369	−0.371
8	24.650	24.650	0.200	0.204)	24.778	24.775	−0.370	−0.371
9	24.842	24.842	0.200	0.201	24.935	24.933	−0.371	−0.375
10	24.978	-	0.199	-	25.049	25.047	−0.372	−0.379
11	25.079	25.082	0.199	0.152	25.134	25.132	−0.372	−0.363
12	25.154	25.158	0.199	0.139	25.199	25.196	−0.373	−0.349
13	25.213		0.198		25.251	25.248	−0.373	−0.364
14	25.259		0.198		25.292		−0.373	
15	25.297		0.198		25.325		−0.373	
16	25.327		0.198		25.352		−0.373	
17	25.352		0.198		25.375		−0.373	
18	25.373		0.198		25.394		−0.373	
19	25.391		0.198		25.410		−0.373	
20	25.406		0.198		25.424		−0.373	
21	25.419		0.198		25.436		−0.373	
22	25.431		0.198		25.446		−0.372	
23	25.440		0.198		25.455		−0.372	
24	25.449		0.198		25.463		−0.372	
25	25.457		0.198		25.470		−0.372	
26	25.463		0.198		25.477		−0.371	
27	25.469		0.198		25.482		−0.371	
28	25.475		0.198		25.487		−0.371	
29	25.479		0.198		25.492		−0.371	
30	25.484		0.198		25.496		−0.370	
31	25.488		0.198		25.500		−0.370	
32	25.491		0.198		25.503		−0.370	
33	25.494		0.198		25.506		−0.370	
34	25.497		0.198		25.509		−0.369	
35	25.500		0.198		25.511		−0.369	
36	25.503		0.198		25.514		−0.369	
37	25.505		0.199		25.516		−0.369	
38	25.507		0.199		25.518		−0.368	
39	25.509		0.199		25.520		−0.368	
40	25.511		0.199		25.522		−0.368	
…								
∞	25.545	25.545			25.555	25.555		

Table 21. Resonance energies (E) and quantum defect (δ) of the $4s^24p^4(^3D_2, {}^1S_0)]ns$, nd series of the Kr^+ ions. The present results from the Modified atomic orbital theory (MAOT) are compared with the Screening constant by unit nuclear charge (SCUNC) results of Sakho [11] are compared with the ALS experimental data of Hinojoha *et al*., [10]. The ALS resonance energies are calibrated to ±30 meV and quantum defects are estimated to within an error of 20%. $\sigma_1(^1D_2) = -0.785 \pm 0.116$; $\sigma_2(^1D_2) = 34.00$; $\sigma_1(^1S_0) = -0.811 \pm 0.116$; $\sigma_2(^1S_0) = 34.00$.

n	$4s^24p^5\,{}^2P^\circ_{1/2} \rightarrow 4s^24p^4(^1S_0)]ns$				$4s^24p^5\,{}^2P^\circ_{3/2} \rightarrow 4s^24p^4(^1S_0)]nd$			
	MAOT	ALS	MAOT	ALS	MAOT	ALS	MAOT	ALS
	E	E	δ	δ	E	E	δ	δ
4	23.738	23.738	0.357	0.358				
5	25.315	25.312	0.358	0.360	25.965	(25.947)	0.369	(0.385)
6	26.132	26.124	0.355	0.368	26.780	26.780	0.380	0.379
7	26.608	26.602	0.353	0.371	27.261	(27.261)	0.381	(0.379)
8	26.910	26.904	0.352	0.376	27.566	27.566	0.381	0.376
9	27.112	27.108	0.351	0.377	27.771	27.771	0.379	0.378
10	27.255	(27.261)	0.351	(0.303)	27.915	27.914	0.378	0.381
11	27.360	27.356	0.351	0.400	28.021	28.020	0.378	0.383
12	27.439	27.435	0.351	0.412	28.100	28.099	0.377	0.381
13	27.500	(27.495)	0.351	(0.435)	28.161	(28.160)	0.377	(0.394)
14	27.548	(27.544)	0.352	(0436)	28.210	(28.208)	0.376	(0.394)
15	27.586		0.352		28.249	(28.247)	0.376	(0.418)
16	27.618		0.353		28.280		0.376	
17	27.644		0.354		28.306		0.376	
18	27.665		0.354		28.328		0.376	
19	27.683		0.355		28.346		0.376	
20	27.699		0.356		28.362		0.376	
21	27.712		0.356		28.375		0.377	
22	27.724		0.357		28.387		0.377	
23	27.734		0.358		28.397		0.377	
24	27.743		0.358		28.405		0.377	
25	27.750		0.359		28.413		0.378	
26	27.757		0.359		28.420		0.378	
27	27.763		0.360		28.426		0.378	
28	27.769		0.361		28.432		0.379	
29	27.774		0.361		28.437		0.379	
30	27.778		0.362		28.441		0.379	
31	27.782		0.362		28.445		0.379	
32	27.786		0.363		28.449		0.380	
33	27.789		0.363		28.452		0.380	
34	27.792		0.364		28.455		0.380	
35	27.795		0.364		28.458		0.381	
36	27.797		0.365		28.460		0.381	
37	27.799		0.365		28.462		0.381	
38	27.802		0.366		28.465		0.382	
39	27.804		0.366		28.467		0.382	
40	27.805		0.366		28.468		0.382	
…								
∞	27.840	27.840			28.503	28.503		

$4s^2 4p^4(^3D_2, {}^1S_0)]ns$, *n*d series of the Kr II ions. Here again, the good agreements between the present MAOT and ALS data [10] may enlighten the accuracy of the uncertain ALS measurement listed into parenthesis. Overall, for both the Cl II, Ar II and Kr II ions, it is demonstrated in this paper the possibilities to reproduce excellently high ALS measurements from single MAOT analytical expression. This is the main strength of the present work.

4. Conclusion

Accurate high lying resonance energy up to $n = 40$ is reported applying the Modified atomic orbital theory. For both the Cl II, Ar II and Kr II ions investigated, a single formula has been established to reproduce with a very good accuracy high experimental measurements such as those performed at the Advanced Light Source at Lawrence Berkeley National Laboratory. A huge number of results are tabulated as useful reference data for interpreting atomic spectra from astrophysical objects containing chlorine, argon and krypton elements.

Conflicts of Interest

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

References

[1] Garnett, D.R. (1989) *The Astrophysical Journal*, **345**, 771.

[2] Küppers, M.E. and Schneider, N.M. (2000) *Geophysical Research Letters*, **25**, 513-516. https://doi.org/10.1029/1999GL010718

[3] Keenan, *et al.* (2003) *The Astrophysical Journal*, **543**, 385.

[4] Hernández, E.M., Juárez, A.M., Kilcoyne, A.L.D., Aguilar, A., Hernández, L., Antillón, A., *et al.* (2015) *Journal of Quantitative Spectroscopy and Radiative Transfer*, **151**, 217-223. https://doi.org/10.1016/j.jqsrt.2014.10.004

[5] McLaughlin, B.M. (2017) *Monthly Notices of the Royal Astronomical Society*, **464,** 1990-1999. https://doi.org/10.1093/mnras/stw2520

[6] Lewis,W. (2005) Chandra Proposal ID# 07500185.

[7] Kraus, M., Borges Fernandes, M., de Araújo, F.X. and Lamers, H.J.G.L.M. (2005) *Astronomy and Astrophysics*, **441**, 289-302. https://doi.org/10.1051/0004-6361:20047013

[8] Kniazev, A.Y., Grebel, E.K., Pustilnik, S.A., Pramskij, A.G. and Zucker, D.B. (2005) *The Astronomical Journal*, **130**,1558-1573. https://doi.org/10.1086/432931

[9] Covington, A.M., Aguilar, A., Covington, I.R., Hinojosa, G., Shirley, C.A., Phaneuf, R.A., *et al.* (2011) *Physical Review A*, **84**, 013413. https://doi.org/10.1103/PhysRevA.84.013413

[10] Hinojoha, G., Covington, A.M., Alna'Washi, G.A., Lu, M., Phaneuf, R.A., Sant'Anna, M.M., *et al.* (2012) *Physical Review A*, **86**, 063402. https://doi.org/10.1103/PhysRevA.86.063402

[11] Sakho, I. (2018) The Screening Constant by Unit Nuclear Charge Method, Description & Application to the Photoionization of Atomic Systems. ISTE Science Publishing Ltd, London and John Wiley & Sons, Inc. USA.

https://doi.org/10.1002/9781119476948

[12] Sakho, I. (2013) *Chinese Journal of Physics*, **51**, 2.

[13] Sakho, I. (2014) *Journal of Atomic and Molecular Science*, **5**, 206.

[14] Sakho, I. (2014) *Journal of Atomic and Molecular Science*, **5**, 206.

[15] Diop, B., *et al.* (2014) *Chinese Journal of Physics*, **52**, 1223.

[16] Sow, M., *et al.* (2014) *Chinese Journal of Physics*, **52**, 1459.

[17] Ralchenko, Y., Kramida, A.E. and Reader, J. (2010) NIST Atomic Spectra Database, (Version 4.0.1). National Institute of Standards and Technology, Gaithersburg.

[18] Ralchenko, Y., Kramida, A.E. and Reader, J. (2011) NIST Atomic Spectra Database (Version 4.0.1). National Institute of Standards, Technology, Gaithersburg, MD.

[19] Ralchenko, Y., Kramida, A.E. and Reader, J. (2014) NIST Atomic Spectra Database (Version 5.2). National Institute of Standards, Technology, Gaithersburg, MD.

Journal of Modern Physics, 2018, 9, 2623-2637
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Reductio ad Absurdum: Modern Physics' Incomplete Absurd Relativistic Mass Interpretation and the Simple Solution That Saves Einstein's Formula

Espen Gaarder Haug

Norwegian University of Life Sciences, Ås, Norway
Email: espenhaug@mac.com

How to cite this paper: Haug, E.G. (2018) Reductio ad Absurdum: Modern Physics' Incomplete Absurd Relativistic Mass Interpretation and the Simple Solution That Saves Einstein's Formula. *Journal of Modern Physics*, **9**, 2623-2637.
https://doi.org/10.4236/jmp.2018.914163

Received: November 12, 2018
Accepted: December 25, 2018
Published: December 28, 2018

Abstract

This paper discusses an absurdity that is rooted in the modern physics' interpretation of Einstein's relativistic mass formula when v is very close to c. Modern physics (and Einstein himself) claimed that the speed of a mass can never reach the speed of light. Yet at the same time they claim that it can approach the speed of light without any upper limit on how close it could get to that special speed. As we will see, this leads to some absurd predictions. If we assert that a material system cannot reach the speed of light, an important question is then, "How close can it get to the speed of light?" Is there a clear-cut boundary on the exact speed limit for an electron, as an example? Or must we settle for a mere approximation?

Keywords

Relativistic Mass, Maximum Velocity of Subatomic Particles, Boundary Condition, Haug Maximum Velocity

1. Introduction

Einstein's relativistic energy mass formula [1] [2] is given by

$$\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{1}$$

Further, Einstein commented on his own formula.

This expression approaches infinity as the velocity v approaches the velocity of light c. The velocity must therefore always remain less than c, however great *may be the energies used to produce the acceleration*[1].

Carmichael (1913) [3] came up with a similar statement in relation to Einstein's theory:

The velocity of light is a maximum which the velocity of a material system may approach but never reach.

We certainly agree with Einstein's formula. Our question is, "How close can v be to c?" Modern physics says nothing about this, except that it can approach c, but never reach c. Does this mean that one can make it as close to c as one wants? This is what we will look into here, and we will show that without a more specific boundary condition on v this can lead to truly absurd predictions.

Einstein's relativistic mass equation predicts that a mass will keep increasing as the velocity of the mass approaches the velocity of the speed of light. If $v = c$, then the mass would become infinite. Einstein and others have given an ad hoc solution to the problem, namely in claiming that indeed the relativistic mass never can become infinite, as this would require an infinite amount of energy for the acceleration. Still, they also seem to claim that the speed of subatomic particles can get as close to c as one would want.

The discussion above is also fully relevant at today's university campus. For example, the excellent text book "University Physics"[2] by Young and Freedman [4] states that

*When the particle's speed v is much less than c, this is approximately equal to the Newtonian expression…**In fact as v approaches c, the momentum approaches infinity.***

Here I have marked part of the sentence in bold. Similarly, in another well-known and excellent university text book by Walker [5] we can read[3].

*As v approaches the speed of light, the relativistic momentum becomes significantly larger than the classical momentum, **eventually diverging to infinity as** $v \to c$.*

Similarly, in the university physics text book by Cutnell and Johnson [6] we can read[4]:

*As v approaches the speed of light c, the $\sqrt{1-v^2/c^2}$ term in the denominator approaches zero. **Hence, the kinetic energy becomes infinitely large**. However, the work-energy theorem tells us that an infinite amount of work would have to be done to give the object an infinite kinetic energy. Since an infinite amount of work is not available, we are left with the conclusion that the objects with mass cannot attain the speed of light c.*

I do not directly disagree; mathematically this is correct. My point is that modern physics does not give an exact limit on how close v can get to c, and we will soon see how this leads to absurd relativistic masses and kinetic energies. In

---

[1] This quote is taken from page 53 in the 1931 edition of Einstein's book *Relativity: The Special and General Theory*. English translation version of Einstein's book by Robert W. Lawson.

[2] 14th edition, page 1238.

[3] Fourth edition, page 1026.

[4] Ninth edition, page 884.

the otherwise excellent book on special relativity by Sartori [7] we can read[5]:

According to equation (7.12), *the kinetic energy of a body approaches infinity when its speed approaches c. This important prediction is confirmed by the experimental data.*

I will claim that these statements are partly wrong, or at least they are not precise. No experiment has shown that the kinetic energy approaches infinity. What has been shown is that the kinetic energy increases rapidly as a particle is accelerated towards a velocity significantly close to the speed of light.

In 1965, Max Born [8] stated that[6]

A glance at Formula (78)[7] *for the mass tells us that the values of the relativistic mass m become greater as the velocity v of the moving body approaches the speed of light. For* $v = c$ *the mass becomes infinitely great. From this it follows that it is impossible to make a body move with a velocity greater than that of light by applying forces: Its inertial resistance grows to an infinite extent and prevents the velocity of light from being reached.*

Long ago, in 1893 Thomson [9] wrote[8]

When in the limit $v = c$ *the increase in mass is infinite, thus the charged sphere moving with velocity of light* ***behaves as if its mass were infinite...***

Naturally, Thomson did not know about Einstein's theory of special relativity, as it was published 12 years later. Still, his equations pointed to a similar result concerning mass when v approaches c.

2. The Absurdity of the Electron Following Modern Physics' Incomplete Relativistic Mass Interpretation

An electron is a very small so-called fundamental particle with a rest-mass of approximately $m_e \approx 9.10938356 \times 10^{-31}$ kg (NIST CODATA 2014). Next let's look at the relativistic mass of the electron as v approaches, but never reaches, the speed of light.

Absurd one Kg mass electron

Assume an electron is accelerated (by a giant exploding star, or by the core of a galaxy, for example) to the following velocity

$$v = c \times 0.99999999999999999999999999999999999999\\ 99999999999999999999999999999999586$$

That is 70 nines behind the decimal point followed by the number 586, or we could say it is 586×10^{-73} with nines instead of zeros after the decimal point. It gives a relativistic mass for a single electron of approximately 1 kg.

Absurd Moon mass electron

Assume an electron is accelerated (by for example a giant exploding star, or by the core of a galaxy) to the following velocity

---

[5] Page 209.

[6] Page 277.

[7] Here Born is referring to the Einstein relativistic mass formula.

[8] Page 21. Actually, Thomson used V as symbol for the speed of light and w for the velocity of the object. We have replaced these with c and v in the citation to make it easier to follow.

$$\begin{aligned} v = c\times{} & 0.999999999999999999999999999999999999999999999 \\ & 9999999999999999999999999999999999999999999999 \\ & 999999999999999999999999923 \end{aligned}$$

That is 116 nines behind the decimal point followed by the number 23, or we could say it is 923×10^{-118} with nines instead of zeros after the decimal point. It gives a relativistic mass for a single electron of approximately 7.34×10^{22} kg, that is basically equal to the rest-mass of the Moon. That is quite amazing, a tiny electron that suddenly has a relativistic mass equal to the rest-mass of the Moon! Where can we find such electrons?

Absurd Earth mass electron

Assume an electron is accelerated to the following velocity

$$\begin{aligned} v = c\times{} & 0.99999999999999999999999999999999999999999999999 \\ & 999999999999999999999999999999999999999999999999 \\ & 999999999999999999999999884 \end{aligned}$$

That is 119 nines behind the decimal point followed by the number 884, or we could say it is 884×10^{-122} with nines instead of zeros after the decimal point. It gives a relativistic mass for a single electron of 5.9806×10^{24} kg, that is basically equivalent to the rest-mass of the Earth.

Absurd Sun mass electron

Assume an electron accelerated to the following velocity

$$\begin{aligned} v = c\times{} & 0.9999999999999999999999999999999999999999999999 \\ & 99999999999999999999999999999999999999999999999 \\ & 9999999999999999999999999999999999999895 \end{aligned}$$

That is 130 nines behind the decimal point followed by the number 895. It gives a relativistic mass for a single electron equal to the rest-mass of the Sun, that is about 1.98×10^{30} kg.

Absurd Milky Way mass electron

Assume an electron is accelerated to the following velocity

$$\begin{aligned} v = c\times{} & 0.9999999999999999999999999999999999999999999999 \\ & 99999999999999999999999999999999999999999999999 \\ & 99999999999999999999999999999999999999999999999 \\ & 999999999999999999999895 \end{aligned}$$

The relativistic mass of the electron at this velocity is equal to the rest-mass of the Milky Way, that is about 10^{12} solar masses. Still, the electron is traveling below the speed of light, so this does not go against mainstream modern physics.

Insane Observable Universe electron

Assume an electron is accelerated to the velocity of

$$\begin{aligned} v = c\times{} & 0.999999999999999999999999999999999999999999999 \\ & 9999999999999999999999999999999999999999999999 \\ & 9999999999999999999999999999999999999999999999 \\ & 99999999999999999999999999999999999996 \end{aligned}$$

That is 174 nines behind the decimal point followed by the number 6, or we could say it is 6×10^{-175} with nines instead of zeros after the decimal point. It gives a relativistic mass for a single electron of approximately 1.018×10^{52} kg, that is basically equal to the rest-mass of what main frame physics claims is the approximate mass of the observable universe, see [10] [11] [12] [13]. That is quite amazing, a tiny electron that suddenly has a relativistic mass equal to the rest-mass of the whole observable universe.

Modern physics leads to absurd kinetic energies for subatomic particles

Table 1 lists the relativistic kinetic energy of an electron traveling at various velocities, all below the speed of light. All of these velocities are valid inside the framework of modern physics, as it stipulates no precise speed limit on the velocity of an electron as long as it falls below the speed of light.

Why don't we see a single electron (or other subatomic particle) with a relativistic mass equal to (even at the most moderate level) the rest-mass of the Moon? Such an electron would have enormous kinetic energy, causing a gigantic impact with collision with the Earth, or other planets in our solar system. We suspect that mainstream physics does not have a good answer to this question. Maybe such fast-traveling electrons exist, but they are rare and therefore have a very low probability of occurring? What if, as a counterpoint, a single electron wiped the dinosaurs out? Are we doomed? And why have we not heard physicists discussing such velocities for electrons? Perhaps they simply do not like to talk about such things, as they have no good explanations for why such very fast electrons have never been observed?

3. A Simple Solution to the Absurdity That Saves Einstein's Relativistic Mass Formula

We have seen how modern physics' assumption that a mass must travel more slowly than the speed of light, while at the same time asserting that it can approach the speed of light, leads to absurd predictions. An important question is, therefore, if there could be an exact speed limit below the speed of light for anything with rest-mass. Haug has recently suggested a maximum velocity for anything with rest-mass by assuming the maximum relativistic mass (energy) for

Table 1. The table shows the kinetic energy for an electron traveling at various velocities below the speed of light.

Velocity of electron % of light (v/c):	Relativistic electron mass = rest-mass of	Kinetic energy:[a]	Ton TNT equivalent:[b]
923×10^{-120} (9'ns in front)	Moon	6.597×10^{39} J	1.58×10^{30}
884×10^{-122} (9'ns in front)	Earth	5.375×10^{41} J	1.28×10^{32}
895×10^{-133} (9'ns in front)	Sun	1.787×10^{47} J	4.27×10^{37}
895×10^{-145} (9'ns in front)	Milky Way	1.787×10^{59} J	4.27×10^{49}

[a] The Kinetic energy is calculated as $E_k=\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}-mc^2$. [b] One ton TNT equivalent is about 4.184 giga joules.

an elementary particle is the Planck mass (energy), see [14]-[19]. The maximum velocity any subatomic particle can take as measured by Einstein-Poincaré synchronized clocks[9] is then given by

$$v_{\max} = c\sqrt{1-\frac{l_p^2}{\bar{\lambda}^2}} \tag{2}$$

where $\bar{\lambda}$ is the reduced Compton wavelength of the mass in question, and l_p is the Planck length [21] [22]. This formula can be derived by setting

$$m_p c^2 \geq \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\sqrt{1-\frac{v^2}{c^2}} \geq \frac{m}{m_p}$$

$$1-\frac{v^2}{c^2} \geq \frac{m^2}{m_p^2}$$

$$v^2 \leq c^2\left(1-\frac{m^2}{m_p^2}\right)$$

$$v \leq c\sqrt{1-\frac{m^2}{m_p^2}} \tag{3}$$

where m_p is the Planck mass, and m is the mass of an elementary particle. And since any elementary mass can be written as

$$m = \frac{\hbar}{\bar{\lambda}}\frac{1}{c} \tag{4}$$

Which means the mass of an elementary particle can be found by measuring the Compton wavelength of the particle, as has been done experimentally with electrons, see [23]. Since the reduced Compton wavelength of the Planck mass must be the Planck length, then we must have

$$v \leq c\sqrt{1-\frac{\left(\frac{\hbar}{\bar{\lambda}}\frac{1}{c}\right)^2}{\left(\frac{\hbar}{l_p}\frac{1}{c}\right)^2}}$$

$$v \leq c\sqrt{1-\frac{l_p^2}{\bar{\lambda}^2}} \tag{5}$$

The same formula can also be derived by assuming that the shortest possible length-contracted Compton wavelength is limited by the Planck length

$$\bar{\lambda}\sqrt{1-\frac{v^2}{c^2}} \geq l_p$$

---

[9] This also holds true if measured with clocks synchronized with very slow clock transportation method, see [20].

$$1-\frac{v^2}{c^2}\geq\frac{l_p^2}{\bar{\lambda}^2}$$

$$v\leq c\sqrt{1-\frac{l_p^2}{\bar{\lambda}^2}} \tag{6}$$

Further, the same maximum velocity of matter can also be found from Heisenberg's uncertainty principle when assuming the uncertainty in position cannot be smaller than the Planck length [24] [25]. So again this would mean the maximum relativistic mass of a elementary particle is the Planck mass. The Planck mass is approximately 2.17651×10^{-8} kg. It is enormous compared to the electron, but it is miniscule compared to the mass of the Moon, Earth, or the Sun. Further, the Planck mass only can last for an instant (the Planck time), as pointed out by Haug, something we soon will get back to soon.

Further, an electron can travel at a velocity very close to that of the speed of light, but its maximum velocity will still be significantly less than what has been described here previously. The maximum velocity for an electron would be approximately

$$v_{\max}=c\sqrt{1-\frac{l_p^2}{\bar{\lambda}_e^2}}\approx c\times0.999999999999999999999999999999999999999999999124 \tag{7}$$

In this calculation, we have assumed the reduced Compton wavelength of the electron given by NIST CODATA, that is $\frac{2.4263102367\times10^{-12}}{2\pi}$ m, and a Planck length of 1.616229×10^{-35} m. Because there is some uncertainty regarding both the exact Planck length and the reduced Compton wavelength, there is some uncertainty around this velocity, but it must be very close to this number. NIST (2014) CODATA reports a standard uncertainty for the Planck length of 0.000038×10^{-35} m. Based on this theory, we can rest assured that the electron (or any other mass) can never reach a relativistic mass even close to one kg, so there is no chance that a single electron will cause much harm (at least not compared to the data in **Table 1**), no matter how fast it is accelerated because there is a maximum velocity that limits both its kinetic energy and its relativistic mass.

Will modern physics accept the existence of a maximum speed limit for subatomic masses or will the field keep holding on to absurd beliefs? If we do not accept the idea that the maximum velocity for subatomic particles has an exact limit below the speed of light, then we must accept the following absurdities:

- That there is a wavelength shorter than the Planck length.
- That there is a time interval shorter than the Planck time.
- That there is a maximum frequency higher than the Planck frequency.
- That an electron can take a relativistic mass similar to that of the Moon, the Earth, the Sun, and even the Milky Way, or even larger masses. This is, at best, truly absurd! Our theory predicts that no subatomic particle can take a relativistic mass higher than the Planck mass.

- That there is no limit on the relativistic Doppler shift. This is also highly unlikely. Haug [15] has suggested that the limit here is the Planck frequency Doppler shift.
- For a subatomic particle, there is a momentum close to infinity. This is absurd. The maximum momentum of a subatomic particle based on our maximum velocity formula for matter is likely just below the Planck momentum.
- For a subatomic particle, there is a kinetic energy close to infinity. This is, again, absurd.

The newly introduced maximum velocity puts a series of limits on subatomic "fundamental particles":

- The maximum frequency is the Planck frequency: $f_{\max}=\frac{c}{l_p}$.
- The maximum relativistic mass a subatomic particle can take is the Planck mass.
- The maximum relativistic momentum a subatomic particle can take is just below the Planck momentum.
- The maximum kinetic energy a subatomic particle can take is close to $\frac{\hbar}{l_p}c$, or more precisely $E_{k,\max}=\hbar c\left(\frac{1}{l_p}-\frac{1}{\bar{\lambda}}\right)$.
- The maximum relativistic length contraction of the reduced Compton wavelength a subatomic particle is l_p, which is the reduced Compton length of the Planck mass particle. This again means the maximum relativistic mass of a elementary particle is the Planck mass.

Also, it is worth mentioning here that the Planck length can be found totally independent of any knowledge of Newton's gravitational constant, see [18] and even independent of any knowledge of the Planck constant, see [26].

4. Ways to Write the Maximum Velocity Formula

There are several ways to write the maximum velocity for subatomic particles that will all give the same answer; here we present some of them.

In terms of reduced Compton wavelength

$$v_{\max}=c\sqrt{1-\frac{l_p^2}{\bar{\lambda}^2}} \tag{8}$$

In terms of particle mass

$$v_{\max}=c\sqrt{1-\frac{m^2}{m_p^2}} \tag{9}$$

where m is the rest-mass of the particle and m_p is the Planck mass.

As a function of Newton's gravitational constant

$$v_{\max}=c\sqrt{1-\frac{Gm^2}{\hbar c}} \tag{10}$$

All of these formulas are basically the same, but each one requires somewhat different input:

$$v_{\max} = c\sqrt{1-\frac{l_p^2}{\bar{\lambda}^2}} = c\sqrt{1-\frac{m^2}{m_p^2}} = c\sqrt{1-\frac{Gm^2}{\hbar c}} \tag{11}$$

The maximum velocity for an electron

For an electron, the maximum velocity can be written as function of the dimensionless gravitational coupling constant.

$$v_{\max} = c\sqrt{1-\alpha_G} \tag{12}$$

This is no surprise, since the dimensionless gravitational coupling constant is given by $\alpha_G = \frac{m_e^2}{m_p^2} = \frac{l_p^2}{\bar{\lambda}_e^2}$. For information about the dimensionless gravitational coupling constant see [27] [28] [29] [30].

5. The Speed of the Planck Mass Particle

In the maximum velocity formula given by Haug, the maximum speed of any observed particle is very close to light, but still faster than what we can achieve at our strongest particle accelerators such as the Large Hadron Collider. However, in the special case where we deal with a Planck mass particle then the reduced Compton wavelength is equal to the Planck length, $\bar{\lambda} = l_p$ and we must have

$$v_{\max} = c\sqrt{1-\frac{l_p^2}{l_p^2}} = 0 \tag{13}$$

That the maximum velocity of the Planck mass particle is zero sounds bizarre at first. However, as suggested by [17] the Planck mass particle is the collision point between two photons. Light always travel at the speed of light, but what is the speed of a photon just at the moment it collides with another photon? We will claim a photon stands still just at the instant (Planck time) it collides with another photon. Recent research has been quite clear on the concept that in a photon-photon collision we likely can create matter, see [31]. We will claim the collision point between two light particles is the missing Planck mass particle. The Planck mass energy $E = p_p c = m_p cc$ seems far too great for any observed photons. However, in our analysis, we claim that the Planck mass particle only can last for one Planck second: $t_p = \frac{l_p}{c}$, before it dissolves into energy and where the light particles once again travel at the speed of light. This means the minimum momentum of a photon is

$$m_p c\frac{l_p}{c} = \frac{\hbar}{c} \tag{14}$$

and multiplied by c we get the energy of the light particle, that is $\hbar$, and dividing this by c^2 we get

$$\frac{\hbar}{c^2} = m_p t_p \tag{15}$$

Haug has recently suggested that all masses ultimately consist of Planck mass particles, but again these Planck mass particle events only last for one Planck second. Schrödinger in 1930 indicated that the electron was in a sort of Zitterbewegung ("trembling motion" in English) $\frac{2mc^2}{\hbar} \approx 1.55269 \times 10^{21}$ per second. We will suggest that the electron is in a Planck mass state $\frac{c}{\bar{\lambda}_e} \approx 7.76344 \times 10^{20}$ per second (about half of that of Schrödinger's "Zitterbewegung" frequency). However, each Planck mass state only lasts for one Planck second and therefore we get the normal electron mass from

$$\frac{c}{\bar{\lambda}_e} m_p \frac{l_p}{c} = \frac{\hbar}{\bar{\lambda}_e} \frac{1}{c} \approx 9.10938 \times 10^{-31} \text{ kg} \tag{16}$$

That is to say, every elementary particle is also a clock ticking at the reduced Compton periodicity. The idea of an internal clock with a clock frequency close to the "Zitterbewegung" frequency in the case of the electron is not new, see [32] [33], for example. The link between mass and Compton time frequency has recently been supported by recent experimental research. Dolce and Perali [34] conclude that "*the rest-mass of a particle is associated to a rest periodicity known as Compton periodicity*". Again, in our model the reduced Compton periodicity is directly linked to a Planck mass event. Between each Planck mass event there is a reduced Compton time interval: $\frac{\bar{\lambda}}{c}$. The Planck mass event itself we suggest only lasts for one Planck second and that even an elementary particle is in an internal energy state most of the time; only in the time fraction $\frac{l_p}{\bar{\lambda}}$ is the particle in a Planck mass state. The rest of the time it is in an internal energy state. An electron switches between energy and mass state approximately 7.76×10^{20} times per second.

This would mean that the mass of an elementary particle such as an electron is time dependent. However, evaluating the time dependency of elementary particles can likely only be done directly when one measures the mass of an elementary particle at time windows close to the reduced Compton time of the particle, where the reduced Compton time simply is defined as the reduced Compton wavelength of the particle divided by the speed of light. This time interval is so short that we are not really able to complete such a study yet, even for an electron.

In 1899, Max Planck was the first to suggest that there was an important elementary mass given by $m_p = \sqrt{\frac{\hbar c}{G}}$. He derived the Planck mass using dimensional analysis, assuming that the Newton gravitational constant, the Planck constant, and the speed of light were the most important universal constants. It was Lloyd Motz, while working at the Rutherford Laboratory in 1962, [35] [36] [37] who first suggested that there likely existed a very fundamental particle with a mass equal to the Planck mass that he called the "Uniton." Motz acknowledged that his Unitons (Planck mass particles) had far too much mass compared to known subatomic masses. The Planck mass one gets from the Max Planck formula is approximately 2.176×10^{-8} kg. Motz tried to explain this by claiming that the Unitons had radiated most of their energy away:

> *According to this point of view, electrons and nucleons are the lowest bound states of two or more Unitons that have collapsed down to the appropriate dimensions gravitationally and radiated away most of their energy in the process.*
> —Lloyd Motz

Others have suggested that there were plenty of Planck mass particles around just after the Big Bang, see [38], but that most of the mass of these super-heavy particles has radiated away. Several physicists, including Motz and Hawking, have suggested that such particles could be micro black holes [39] [40] [41]. Planck mass particles have even been proposed as candidates for cosmological dark matter [42] [43].

Still, no sign of the enormous Planck mass particle has been found, and there have been no observations of micro black holes yet either. The Compton clock model of matter seems to give a possible simple solution to this puzzle. In our model of matter, one needs to consider the observational time window, in particular when it comes to the Planck mass particle. The Planck mass particle likely only lasts for one Planck second. So, if observed inside one Planck second its mass is indeed its well-known Planck mass, and it is then very large compared to any observed elementary particles so far. However, we are not even close to observing particles at the Planck time scale. In a longer time window, one second for example, the Planck mass particle is only approximately 1.173×10^{-51} kg. That is, we have likely looked for a much too large mass in our search for the Planck mass particle and micro black holes. With that in mind, we should perhaps change the perspective on what a micro black hole is. In our view, it is nothing more than the very collision point between two photons: they collide and the collision lasts for one Planck second. At collision the photons stand absolutely still, and then dissolve into energy again moving at the speed of light. The micro black hole is, in this way, almost a misnomer. Further theoretical and practical research is needed to decide if this is the case or not. In particular, research looking for possible breaks in Lorentz symmetry could be important here and is what we come to in the next section.

6. Breakdown of Lorentz Invariance at the Planck Scale?

The maximum velocity formula for anything with rest-mass would mean Lorentz invariance breaks down at the Planck scale. Based on this view, the Planck particle, the Planck length, and the Planck time, unlike any other particle, length, or time, seem to be the same no matter what frame they are observed from. The view that Lorentz invariance could be broken at the Planck scale appears to be consistent with what is predicted by several quantum gravity theories, see for example [44]. Lorentz symmetry is supported by a long series of tests, but it has never been tested at anything even close to the Planck scale (at distances close to the Planck length, or Planck energies), so one should be careful to use experimental evidence as an argument against this idea.

One could ask to what degree new physics at the Planck scale could be weakly detected at lower energies; this is discussed by [45] [46], for example. A recent review article [47] on the possibility for Lorentz symmetry breaking in relation to quantum gravity predictions and experiments noted:

In conclusion, though no violation of Lorentz symmetry has been observed so far, an incredible number of opportunities still exist for additional investigations.

So, we think testable predictions related to the Planck scale should be investigated further, also in relation to our maximum velocity formula. Recently, Haug [48] has indicated that this new maximum velocity of matter likely predicts zero velocity time-dilation in quasars, which is consistent with what has been observed [49] [50]. The Schwarzschild radius is linked to the Planck length, $R_s = 2\frac{Gm}{c^2} = 2N\frac{l_p^2}{\bar{\lambda}}$, where N is the number of Planck masses in the mass in question. This, combined with the fact that quasars are often considered black holes (at least at their core), means quasars likely are linked to the Planck scale. It could be that we already have observations from the Planck scale from quasars, and that such anomalies as lack of time dilation in quasars can distinguish various theories in relation to their prediction of how Lorentz symmetry is broken down or not broken down. At least it seems that our theory in relation to this maximum velocity likely is consistent with everything that has been observed so far.

7. Conclusion

We conclude that in stating that a mass must travel more slowly than the speed of light, while at the same time asserting that it can approach the speed of light, we get absurd predictions. Examples include the idea that an electron could attain a relativistic mass equal to the rest-mass of the Moon, the Earth, the Sun, the Milky Way, or even entire galaxy clusters. Haug has recently addressed this absurdity by showing that there must be a precise maximum velocity for anything with mass likely given by $v_{\max} = c\sqrt{1-\frac{l_p^2}{\bar{\lambda}^2}}$.

Acknowledgements

Thanks to Victoria Terces for helping me edit this manuscript. Also thanks to Alan Lewis, Daniel Duffy, ppauper, and AvT for useful tips on how to do high precision calculations. Also thanks to an anonymous referee for very useful comments.

Conflicts of Interest

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

References

[1] Einstein, A. (1905) *Annalen der Physik*, **323**, 639-641. https://doi.org/10.1002/andp.19053231314

[2] Einstein, A. (1916) Relativity: The Special and the General Theory. Translation by Robert Lawson (1931), Crown Publishers, New York.

[3] Carmichael, R.D. (1913) The Theory of Relativity. John Wiley & Sons, Hoboken.

[4] Freedman, R.A. and Young, H.D. (2016) University Physics with Modern Physics, 14th Edition, Pearson, Boston.

[5] Walker, J.S. (2010) Physics. 4th Edition, Addison-Wesley, Boston.

[6] Cutnell, J.D. and Johnson, K.W. (2012) Physics. 9th Edition, John Wiley & Sons, Inc., Hoboken.

[7] Sartori, L. (1996) Understanding Relativity. University of California Press, Oakland.

[8] Born, M. (1965) Einstein's Theory of Relativity. Dover, Mineola, New York.

[9] Thomson, J.J. (1893) Recent Researches in Electricity and Magnetism. Oxford University Press, Oxford.

[10] Whitrow, G.J. (1946) *Nature*, **185**, 165-166. https://doi.org/10.1038/158165b0

[11] Dirac, P. (1947) The Principles of Quantum Mechanics. Oxford University Press, Oxford.

[12] Carvalho, J.C. (1995) *International Journal of Theoretical Physics*, **34**, 2507-2509. https://doi.org/10.1007/BF00670782

[13] Persinger, M.A. (2009) *Journal of Physics*, *Astrophysics and Physical Cosmology*, **3**, 1-3.

[14] Haug, E.G. (2016) The Planck Mass Particle Finally Discovered! Good Bye to the Point Particle Hypothesis! http://vixra.org/pdf/1607.0496v7.pdf

[15] Haug, E.G. (2016) A New Solution to Einstein's Relativistic Mass Challenge Based on Maximum Frequency. http://vixra.org/pdf/1609.0083v2.pdf

[16] Haug, E.G. (2016) *Physics Essays,* **29***,* 558-561.

[17] Haug, E.G. (2017) *Acta Astronautica*, **136**, 144-147. https://doi.org/10.1016/j.actaastro.2017.03.011

[18] Haug, E.G. (2017) *Applied Physics Research*, **9**, 58.

[19] Haug, E.G. (2018) *International Journal of Astronomy and Astrophysics*, **8**, 6-23. https://doi.org/10.4236/ijaa.2018.81002

[20] Haug, E.G. (2014) Unified Revolution: New Fundamental Physics. E.G.H. Publishing, Oslo.

[21] Planck, M. (1899) Naturlische Masseinheiten. Der Königlich Preussischen Akademie Der Wissenschaften, Berlin, 479.

[22] Planck, M. (1906) Vorlesungen über die Theorie der Wärmestrahlung. J.A. Barth, Leipzig, 163.

[23] Prasannakumar, S., Krishnaveni, S. and Umesh, T.K. (2012) *European Journal of Physics*, **33**, 65-72. https://doi.org/10.1088/0143-0807/33/1/005

[24] Haug, E.G. (2018) Does Heisenberg's Uncertainty Principle Predict a Maximum Velocity for Anything with Rest-Mass below the Speed of Light? http://vixra.org/pdf/1801.0218v4.pdf

[25] Haug, E.G. (2018) Revisiting the Derivation of Heisenberg's Uncertainty Principle: The Collapse of Uncertainty at the Planck Scale. https://doi.org/10.20944/preprints201805.0258.v2

[26] Haug, E.G. (2018) Finding the Planck Length Independent of Newton's Gravitational Constant and the Planck Constant: The Compton Clock Model of Matter. https://www.preprints.org/manuscript/201809.0396/v1

[27] Silk, J. (1977) *Nature*, **265**, 710-711. https://doi.org/10.1038/265710a0

[28] Rozental, I.L. (1980) *Soviet Journal of Experimental and Theoretical Physics Letters,* **31**, 19-27.

[29] Neto, M.D.O. (2005) *Chaos, Solitons and Fractals*, **24**, 19-27. https://doi.org/10.1016/j.chaos.2004.07.032

[30] Burrows, A.S. and Ostriker, J.P. (2013) Astronomical Reach of Fundamental Physics. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 31-36.

[31] Pike, O.J., Mackenroth, F., Hill, E.G. and Rose, S.J. (2014) *Nature Photonics*, **8**, 434-436. https://doi.org/10.1038/nphoton.2014.95

[32] Rivas, M. (2008) Measuring the Internal Clock of the Electron.

[33] Bauer, M. (2014) Electron Channeling, de Broglie's Clock and the Relativistic Time Operator.

[34] Dolce, D. and Perali, A. (2015) *The European Physical Journal Plus*, **130**, 41. https://doi.org/10.1140/epjp/i2015-15041-5

[35] Motz, L. (1962) *Il Nuovo Cimento*, **26**, 672-697. https://doi.org/10.1007/BF02781795

[36] Motz, L. (1966) A Gravitational Theory of the mu meson and Leptons in General. Rutherford Observatory, Columbia University, New York.

[37] Motz, L. (1971) The Quantization of Mass. Rutherford Observatory, Columbia University, New York.

[38] De, S.S. (2001) *International Journal of Theoretical Physics*, **40**, 2067-2084. https://doi.org/10.1023/A:1011953623319

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

[40] Motz, L. and Epstein, J. (1979) *Il Nuovo Cimento*, **51**, 88-113. https://doi.org/10.1007/BF02822327

[41] Obermair, G.M. (2013) *Journal of Physics, Conference Series*, **442**, Article ID: 012066. https://doi.org/10.1088/1742-6596/442/1/012066

[42] Macgibbon, J.H. (1987) *Nature*, **329**, 308-309. https://doi.org/10.1038/329308a0

[43] Dokuchaev, V.I. and Eroshenko, Y.N. (2014) Black Hole Atom as a Dark Matter Particle Candidate. https://arxiv.org/abs/1403.1375 https://doi.org/10.1155/2014/434539

[44] Kislat, F. and Krawczynski, H. (2017) *Physical Review D*, **95**, Article ID: 083013. https://doi.org/10.1103/PhysRevD.95.083013

[45] Amelino-Camelia, G., Ellis, B.J., Mavromatos, N.E., Nanopoulos, D.V. and Sarkar, S. (2998) *Nature*, **393**, 763-765. https://doi.org/10.1038/31647

[46] Reyes, C.M., Ossandon, S. and Reyes, C. (2005) *Physics Letters B*, **746**, 190-193. https://doi.org/10.1016/j.physletb.2015.05.006

[47] Hees, A., *et al.* (2017) *Universe*, **2**, 30. https://doi.org/10.3390/universe2040030

[48] Haug, E.G. (2018) Maximum Velocity for Matter in Relation to the Schwarzschild Radius Predicts Zero-Time Dilation for Quasars.

[49] Hawkings, M.R.S. (2000) *Monthly Notices of the Royal Astronomical Society*, **406**, 1047-1054.

[50] Hawkings, M.R.S. (2001) *The Astrophysical Journal*, **553**, 97-100. https://doi.org/10.1086/320683

Journal of Modern Physics, 2018, 9, 2638-2656
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Pseudoscalar Top-Bottom Quark-Antiquark Composite as the Resonance with 28 GeV at the LHC: Hadron Masses and Higgs Boson Masses Based on the Periodic Table of Elementary Particles

Ding-Yu Chung

Utica, Michigan, USA
Email: dy_chung@yahoo.com

How to cite this paper: Chung, D.-Y. (2018) Pseudoscalar Top-Bottom Quark-Antiquark Composite as the Resonance with 28 GeV at the LHC: Hadron Masses and Higgs Boson Masses Based on the Periodic Table of Elementary Particles. *Journal of Modern Physics*, **9**, 2638-2656.
https://doi.org/10.4236/jmp.2018.914164

Received: December 1, 2018
Accepted: December 25, 2018
Published: December 28, 2018

Abstract

This paper posits that the observed resonance with 28 GeV at the LHC is the pseudoscalar top-bottom quark-antiquark composite which has the calculated mass of 27.9 GeV derived from the periodic table of elementary particles. The calculated mass is for the mass of $b\bar{b} + (b\bar{b} + t\bar{t})/2$. In the periodic table of elementary particles, t quark (13.2 GeV) in the pseudoscalar top-bottom quark-antiquark composite is only a part of full t quark (175.4 GeV), so pseudoscalar $t\bar{t}$ (26.4 GeV) cannot exist independently, and can exist only in the top-bottom quark-antiquark composite. As shown in the observation at the LHC, the resonance with 28 GeV weakens significantly at the higher energy collision (13 TeV), because at the higher collision energy, low-mass pseudoscalar $t\bar{t}$ in the composite likely becomes independent full high-mass vector $t\bar{t}$ moving out of the composite. The periodic table of elementary particles is based on the seven mass dimensional orbitals derived from the seven extra dimensions of 11 spacetime dimensional membrane. The calculated masses of hadrons are in excellent agreement with the observed masses of hadrons by using only five known constants. For examples, the calculated masses of proton, neutron, pion ($\pi^{\pm}$), and pion (π^{0}) are 938.261, 939.425, 139.540, and 134.982 MeV in excellent agreement with the observed 938.272, 939.565, 139.570, and 134.977MeV, respectively with 0.0006%, 0.01%, 0.02%, and 0.004%, respectively for the difference between the calculated and observed mass. The calculated masses of the Higgs bosons as the intermediate vector boson composites are in excellent agreements with the observed masses. In conclusion, the calculated masses of the top-bottom quark-antiquark composite (27.9 GeV), hadrons, and the Higgs bosons by the periodic table of elementary particles are in excellent agreement with the observed masses of resonance with 28 GeV at the LHC, hadrons, and the Higgs bosons, respectively.

Keywords

LHC, CMS, Resonance, b Quark Jet, Periodic Table of Elementary Particles, Top Quark, Bottom Quark, Hadron Masses, Mass Calculation, Higgs Boson

1. Introduction

In the search for resonances in the mass range 12 - 70 GeV produced in association with a b quark jet and a second jet, and decaying to a muon pair, the CMS Collaboration at the LHC recently reported an excess of events above the background near a dimuon mass of 28 GeV [1]. The search is carried out in two mutually exclusive event categories from proton-proton collisions at center-of-mass energies of 8 and 13 TeV. The first category involves a b quark jet in the central region and at least one jet in the forward region, while the second category involves two jets in the central region, at least one of which is identified as a b quark jet, no jets in the forward region. At the 8 TeV collision, the first category has 4.2 standard deviations, while the second category has 2.9 standard deviations. At the 13 TeV collision, the first category has 2.0 standard deviations, while the second category results in a 1.4 standard deviation deficit.

This potential new particle at 28 GeV does not match the properties of any of particles in the standard model. It is also puzzling that the resonance at 28 GeV weakens, disappears, or gets inverted at 13 TeV. This paper posits that the resonance with 28 GeV observed recently at the LHC is the pseudoscalar top-bottom quark-antiquark composite which has the calculated mass of 27.9 GeV derived from the periodic table of elementary particles in good agreement with the observed 28 GeV. The calculated mass is the mass of three pseudoscalar b quarks and one pseudoscalar t quark which represent the composite of $b_p\bar{b}_p$ + $(b_p\bar{b}_p + t_p\bar{t}_p)/2$ where p = pseudoscalar. (The quark in pseudoscalar meson is denoted as "pseudoscalar quark", while the quark in vector mesons is denoted as "vector quarks" which has higher mass than pseudoscalar quark.) As described in the periodic table of elementary particles, pseudoscalar t quark (13.2 GeV) is only a part of full t quark (175.4 GeV), so pseudoscalar $t_p\bar{t}_p$ (26.4 GeV) cannot exist independently, and can exist only in the top-bottom quark-antiquark composite. As shown in the observation at the LHC, the resonance with 28 GeV weakens significantly at the higher energy collision (13 TeV), because at the higher collision energy, low-mass pseudoscalar $t\bar{t}$ in the composite likely becomes independent full high-mass vector $t\bar{t}$ moving out of the composite. To account for the observed two jets, the composite has two jets consisting of a $b\bar{b}$ jet and a b + t jet, where $b\bar{b}$ jet for $(b_p\bar{b}_p + t_p\bar{t}_p)/2$ is more stable than b + t jet which decays faster into the jet in the forward region to constitute the first category of the search by the CMS Collaboration at the LHC.

The periodic table of elementary particles is based on the seven mass dimensional orbitals derived from the seven extra dimensions of 11 spacetime dimensional membrane particles [2] [3] [4]. The seven mass dimensional orbitals include the seven principal mass dimensional orbitals for stable baryonic matter leptons (electron and neutrinos), gauge bosons, gravity, and dark matter and the seven auxiliary mass dimensional orbitals for unstable leptons (muon and tau) and quarks, and calculate accurately the masses of all elementary particles and the cosmic rays by using only five known constants [5] [6]. Hadron masses can be calculated in excellent agreement with the observed masses of hadrons. For examples, the calculated masses of proton, neutron, pion ($\pi^{\pm}$), and pion (π^{0}) are 938.261, 939.425, 139.540, and 134.982 MeV in excellent agreement with the observed 938.272, 939.565, 139.570, and 134.977 MeV, respectively with 0.0006%, 0.01%, 0.02%, and 0.004%, respectively for the difference between the calculated and observed mass.

Section 2 describes the periodic table of elementary particles and the mass formulas. Section 3 deals with quarks and hadrons. Section 4 explains the top-bottom quark-anti-quark composite as the resonance at 28 GeV. Section 5 describes the Higgs boson doublet as the intermediate vector boson composites.

2. The Periodic Table of Elementary Particles and the Mass Formulas

The periodic table of elementary particles [2] [3] [4] is based on the seven mass dimensional orbitals derived from the seven extra dimensions of 11 spacetime dimensional membrane. The seven mass dimensional orbitals include the seven principal mass dimensional orbitals for stable baryonic matter leptons (electron and neutrinos), gauge bosons (all forces), gravity, and dark matter (five sterile dark matter neutrinos) and the seven auxiliary mass dimensional orbitals for unstable leptons (muon and tau) and quarks (d, u, s, c, b, and t) as in **Figure 1** and **Table 1**.

Table 1. The periodic table of elementary particles for baryonic matter and dark matter.

d	a = 0	a = 0	1	2	1	2	3	4	5	a = 0
	Stable Baryonic Matter Leptons	Dark Matter Leptons	Unstable Leptons		Quarks					Bosons
5	ν_e	ν_{DM5}								B_5 = A electromagnetism
6	e	ν_{DM6}								B_6 = g* strong (basic gluon)
7	ν_μ	ν_{DM7}	μ_7	τ_7	d_7/u_7	s_7	c_7	b_7	t_7	$B_7 = Z^0_L$ left-handed weak
8	ν_τ	ν_{DM8}	μ_8 (absent)		b_8 (absent)	t_8				$B_8 = X_R$ right-handed CP
9	ν'_τ (high-mass ν_τ)	ν_{DM9}								$B_9 = X_L$ left-handed CP
10										$B_{10} = Z^0_R$ right-handed weak
11										B_{11} = gravity

d = principal mass dimensional orbital number, a = auxiliary mass dimensional orbital number.

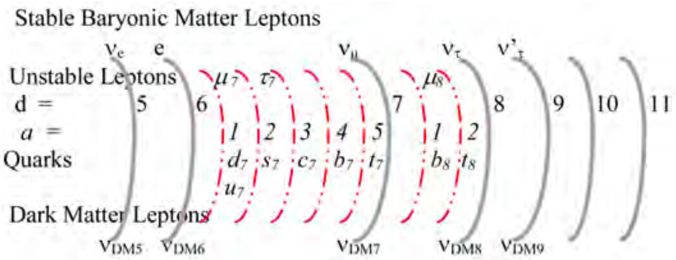


Figure 1. The seven principal mass dimensional orbitals (solid lines) denoted by the principal mass dimensional orbital number d and the seven auxiliary mass dimensional orbitals (dash-dotted lines) denoted by the auxiliary mass dimensional orbital number *a*.

The periodic table of elementary particles calculates accurately the particle masses of all leptons, quarks, gauge bosons, the Higgs boson, and the cosmic rays by using only five known constants: the number (seven) of the extra spatial dimensions in the observed four-dimensional spacetime from the eleven-dimensional membrane, the mass of electron, the masses of Z and W bosons, and the fine structure constant [5] [6]. The calculated masses are in excellent agreements with the observed masses. For examples, the calculated masses of muon, top quark, pion, neutron, and the standard model Higgs boson are 105.55 MeV, 175.4 GeV, 139.54 MeV, 939.42 MeV, and 126 GeV, respectively, in excellent agreements with the observed 105.65 MeV, 172.4 GeV, 139.57 MeV, 939.27 MeV, and 126 GeV, respectively.

The seven mass dimensional orbitals are arranged as F_5 B_5 F_6 B_6 F_7 B_7 F_8 B_8 F_9 B_9 F_{10} B_{10} F_{11} B_{11}, where F_d and B_d are mass dimensional fermion and mass dimensional boson, respectively. As described in the previous papers [2] [3], the mass of mass dimensional fermion and the mass of mass dimensional boson are related to each other with three simple formulas as the follows.

$$M_{d,B} = M_{d,F} / \alpha_d \tag{1}$$

$$M_{d+1,F} = M_{d,B} / \alpha_{d+1} \tag{2}$$

$$M_{d+1,B} = M_{d,B} / \alpha_{d+1}^2 \tag{3}$$

where d is the mass dimensional orbital number, F is fermion, and B is boson. Each dimension has its own α_d, and all α_d's except α_7 (α_w) of the seventh dimension (weak interaction) are equal to α, the fine structure constant of electromagnetism. The given observed masses are the mass of electron for F_6 and the mass of Z boson for B_7. From Equations (1) and (3), $\alpha_w = \alpha_7 = \alpha$ of week interaction = $(M_{B6}/M_{B7})^{1/2} = (M_{F6}/\alpha/M_{B7})^{1/2} = (M_e/\alpha/M_Z)^{1/2} = 0.02771$. Therefore, the masses of gauge bosons are as in Table 2.

The lowest energy gauge boson (B_5) at d = 5 is the Coulomb field for electromagnetism. The second gauge lowest boson (B_6) at d = 6 is basic gluon (g* = 70 MeV ≈ one half of pion) is the strong force as the nuclear force in the pion theory [7] where pions mediate the strong interaction at long enough distances (longer than the nucleon radius) or low enough energies. At short enough distances (shorter than the nucleon radius) or high enough energies, gluons emerge

Table 2. The masses of the principal mass dimensional orbitals (gauge bosons).

B_d	M_d	GeV (calculated)	Gauge boson	Interaction
B_5	$M_e\alpha$	3.7×10^{-6}	A = photon	Electromagnetic
B_6	M_e/α	7×10^{-2} (70.02 MeV)	g* = basic gluon	Strong
B_7	$M_Z = M_{B6}/\alpha_w^2$	91.1876 (given)	Z_L^0	weak (left)
B_8	$M_7/\alpha^2 = M_Z/\alpha^2$	1.71×10^6	X_R	CP (right) nonconservation
B_9	$M_8/\alpha^2 = M_Z/\alpha^4$	3.22×10^{10}	X_L	CP (left) nonconservation
B_{10}	$M_9/\alpha^2 = M_Z/\alpha^6$	6.04×10^{14}	Z_R^0	weak (right)
B_{11}	$M_{10}/\alpha^2 = M_Z/\alpha^8$	1.13×10^{19}	G	Gravity

to confine fractional charge quarks. B_6 is denoted as basic gluon, g*. The third lowest boson (B_7) at d = 7 is Z^0 for the weak interaction.

F_{11} (8.275×10^{16} GeV) relates to spin 3/2 gravitino, while B_{11} (1.134×10^{19} GeV) relates to spin 2 graviton. In supersymmetry, gravitino and graviton mediate the supersymmetry between fermion and boson in space dimension and gravitation. There are 10 space dimensions in the 11 spacetime dimensional membrane. As a result, the supersymmetry involves $10F_{11} + B_{11}$, which is equal to 1.217×10^{19} GeV in excellent agreement with the Planck mass (1.221×10^{19} GeV) derived from observed gravity as $(\hbar c/G)^{1/2}$ where c is the speed of light, G is the gravitational constant, and ħ is the reduced Planck constant.

The lepton mass formula and the quark mass formula are derived from the incorporation of basic gluon (g* = B_6 = 70 MeV) to electron. The incorporation of basic gluon as flux quanta follows the the composite fermion theory for the FQHE (fractional quantum Hall effect) [8] [9]. In the composite fermion model for FQHE, the formation of composite fermion is through the attachment of an even number of magnetic flux quanta to electron, while the formation of composite boson is through the attachment of an odd number of magnetic flux quanta to electron. In the same way, the formation of composite fermion is through the attachment of an even number of basic gluons to electron, while the formation of composite boson is through the attachment of an odd number of basic gluons to electron. The formation of composite boson is equal to the formation of composite di-leptons, so the formation of composite lepton is through the attachment of one half of an odd number of basic gluons to electron. As a result, the muon (μ) mass formula is as follows.

$$\begin{aligned} M_{\mu_7} &= Me + 3M_{g^*}/2 \\ &= Me + 3M_e/2\alpha \\ &= 105.5488\ \text{MeV} \end{aligned} \tag{4}$$

which is in excellent agreement with the observed 105.6584 MeV [10] for the mass of muon. The mass of τ follows the Barut lepton mass formula [11] as follows.

$$M_{\text{lepton}} = M_e + \frac{3M_e}{2\alpha}\sum_{a=0}^{n} a^4 \tag{5}$$

where $a = 0$, 1, and 2 are for e, μ_7, and τ_7, respectively. The calculated mass of τ_7 is 1786.2 MeV in good agreement with the observed mass as 1776.82 MeV. According to Barut, the second term, $\sum_{a=0}^{n} a^4$ of the mass formula is for the Bohr-Sommerfeld quantization for a charge-dipole interaction in a circular orbit. The more precise calculated mass of τfor the tau lepton mass formula is as follows.

$$\begin{aligned} M_\tau &= Me + \left(\frac{3M_e}{2\alpha} - M_e\right)\sum 2^4 \\ &= Me + \left(17\frac{3M_e}{2\alpha} - 17M_e\right) \\ &= 1777.47\ \text{MeV} \end{aligned} \tag{6}$$

which is in excellent agreement with observed 1776.82 MeV, and means that during this dipole-interaction in a circular orbit for τ, an electron with total mass of 17M$_e$ is lost. 17M$_e$ is shown as the observed 17 MeV for 34M$_e$ in the light boson (17 e$\overline{\text{e}}$) [12] [13].

Quark has fractional charge (±1/3 or ±2/3), 3-color gluons (red, green, and blue) for 3g*, and both the principal mass dimensional orbitals and axillary mass dimensional orbitals, so similar to Equation (4), d and u in the principal mass dimensional orbital involves e/3 or 2e/3 and 3g* as follows.

principal mass dimensional orbital at d = 6

$$\begin{aligned} M_{\text{principal } q} &= \frac{1 \text{ or } 2Me}{3} + \frac{3\left(3M_{\text{g*}}\right)}{2} \\ &= \frac{1 \text{ or } 2Me}{3} + \frac{3\left(3M_{\text{B6}}\right)}{2} \\ &= \frac{1 \text{ or } 2Me}{3} + \frac{9Me}{2\alpha} \end{aligned} \tag{7}$$

For quarks in the auxiliary mass dimensional orbitals, 3-color basic gluons (3g*) become 3-color auxiliary basic gluons (3g^*_{a7}) at d = 7. Based on Equation (2), auxiliary basic gluon is derived from muon as follows.

$$M_{3\text{g}^*_{a7}} = M_{\mu_7}\alpha_w \tag{8}$$

Similar to Equation (5), the masses of quarks in the auxiliary mass dimensional orbital are as follows.

auxiliary mass dimensional orbital at d = 7

$$M_{\text{auxiliary } q7} = \frac{3\left(3M_{g^*_{a7}}\right)}{2}\sum_{a=1}^{n} a^4 = \frac{9M_{\mu_7}\alpha_w}{2}\sum_{a=1}^{n} a^4 \tag{9}$$

The quark mass formula at d = 7 is the combination of Equations (7) and (9) as follows.

$$M_{q7} = \frac{1 \text{ or } 2Me}{3} + \frac{9Me}{2\alpha} + \frac{9M_{\mu_7}\alpha_w}{2}\sum_{a=1}^{n} a^4 \tag{10}$$

where a = 1, 2, 3, 4, and 5 for u_7/d_7, s_7, c_7, b_7, and t_7, respectively.

The quark mass at a = 5 for the auxiliary mass dimensional orbital at d = 7 is the maximum mass below the mass of B_7, so the next auxiliary mass dimensional orbital has to start from B_7.There are b and t at d = 8, so it is necessary to have $\mu_8$for the masses of b and t. Like μ_7 in Equation (4), the mass of μ_8 is as follows.

$$\begin{aligned} M_{\mu_8^0} &= 2Me + 3M_{g_7^*}/2 \\ &= 2Me + 3M_{B_7}/2 \\ &= 2Me + 3M_{Z^0}/2 \\ &= 136.78\ \text{GeV} \end{aligned} \tag{11}$$

Since at d = 7, there are 3-color basic gluons, at d = 8, 3-*color* basic gluons are not needed, and only one basic gluon (g_7^*) at d = 7 is used. Similar to Equations (7) and (9). The quark mass formulas for the principal and auxiliary mass dimensional orbitals are as follows.

$$\begin{aligned} &\text{principal mass dimensional orbital at d} = 7 \\ &M_{\text{principal quark}} = 3M_{g_7^*}/2 = 3M_{B_7}/2 = 3M_Z/2 \end{aligned} \tag{12}$$

$$\begin{aligned} &\text{auxiliary mass dimensional orbital at d} = 8 \\ &M_{\text{auxiliary quark}} = \frac{3\left(M_{g_{a8}^*}\right)}{2}\sum_{a'=1}^{n'} a'^4 = \frac{3\mu_8^0\alpha}{2}\sum_{a'=1}^{n'} a'^4 \end{aligned} \tag{13}$$

The quark mass formula at d = 8 is the combination of Equations (12) and (13) as follows.

$$M_{q8} = \frac{3M_Z}{2} + \frac{3M_{\mu_8^0}\alpha}{2}\sum_{a'=1}^{n'} a'^4 \tag{14}$$

where a'= 1 and 2 for b_8 and t_8, respectively.

Combining Equations (10) and (14), the quark mass formula is as follows.

$$M_{\text{quark}} = \frac{1\text{ or }2Me}{3} + \frac{9Me}{2\alpha} + \frac{9M_{\mu_7}\alpha_w}{2}\sum_{a=1}^{n} a^4 + \frac{3M_Z}{2} + \frac{3M_{\mu_8^0}\alpha}{2}\sum_{a'=1}^{n'} a'^4 \tag{15}$$

where a =1, 2, 3, 4, and 5 for d/u. s, c, b, and t, respectively, and a'= 1 and 2 for b and t respectively. The calculated masses for d, u, s, c, b, and t are 328.4 MeV, 328.6 MeV, 539 MeV, 1605.3 MeV, 4974.6 MeV, and 175.4 GeV, respectively. In the standard model, there are three generations of leptons. Extra-muon μ_8 is outside of the three generations of leptons in the standard model, so μ_8 is absent as shown in Table 2. As shown in Table 2, to be symmetrical to the absent μ_8, b_8 quark is also absent. The calculated mass of top quark is 175.4 GeV in good agreement with the observed 172.4 GeV. The calculated masses are comparable to the quark masses proposed by De Rujula, Georgi, and Glashow [14], Griffiths [15], and El Naschie [16].

3. Quarks and Hadrons

For baryons, the quarks in the periodic table of elementary particles are baryonic quarks. Mesons have vector mesons with parallel spins and pseudoscalar mesons with antiparallel spins. Since parallel spins have higher energy than antiparallel spins, vector mesons have higher masses than pseudoscalar mesons. For higher-mass vector mesons (parallel spins), vector t, b, c, and s are baryonic t, b, c, s, and vector d and u are baryonic d and u plus basic gluon (g* = 70 MeV). For lower-mass pseudoscalar mesons (antiparallel spins), pseudoscalar t as t_7 is a part of vector t (= t_7 + t_8), pseudoscalar b and c are vector quarks minus g*, and pseudoscalar d, u, and s are derived from g* as shown in **Table 3**.

The calculated masses and the observed masses [10] of baryons are listed in **Table 4**. The binding energy for each d or u quark involves the auxiliary mass dimensional orbital at d = 7 from Equation (9). The primary binding energy E_{Q1} for d or u quark from Equation (9) is as follows.

$$E_{Q1} = 9M_{\mu_7}\alpha_w/2 = 13.162\ \text{MeV} \tag{16}$$

Table 3. The masses of quarks.

Particle	Symbol	Composition	d=	a=	Charge	Generation	Mass (calculated)
Electron	e	e	6	0	−1	1	0.511 MeV (given)
Basic gluon	g*	B_6	6	0	0		70.02 MeV
Baryon							
d	d_b	d_7	7	1	−1/3	1	328.5 MeV
u	u_b	u_7	7	1	2/3	1	328.6 MeV
s	s_b	s_7	7	2	−1/3	2	539.0 MeV
c	c_b	c_7	7	3	2/3	2	1605.3 MeV
b	b_b	b_7	7	4	−1/3	3	4974.7 MeV
t	t_b	$t_7 + t_8$	7 + 8	5 + 2	2/3	3	175.4 GeV
Vector meson							
d	d_v	d_7 + g*	6 + 7	0 + 1	−1/3	1	398.5 MeV
u	u_v	u_7 + g*	6 + 7	0 + 1	2/3	1	398.6 MeV
s	s_v	s_7	7	2	−1/3	2	539.0 MeV
c	c_v	c_7	7	3	2/3	2	1605.3 MeV
b	b_v	b_7	7	4	−1/3	3	4974.7 MeV
t	t_v	$t_7 + t_8$	7 + 8	5 + 2	2/3	3	175.4 GeV
Pseudoscalar meson							
d	d_p	g* + 1/3e	6	0	−1/3	1	70.2 MeV
u	u_p	g* + 2/3e	6	0	2/3	1	70.4 MeV
s	s_p	2(3g* + 3e) + 1/3e	6	0	−1/3	2	423.4 MeV
c	c_p	c_7 − g*	6 + 7	0 + 3	2/3	2	1535.3 MeV
b	b_p	b_7 − g*	6 + 7	0 + 4	−1/3	3	4904.7 MeV
t	t_p	t_7	7	5	2/3	3	13.2 GeV

d = principal mass dimensional mass orbital number, a = auxiliary mass dimensional number, generation = generation of lepton-quark in the standard model.

Table 4. The masses of baryons.

Baryon	Composition	Calculated mass without binding energy MeV	Calculated mass with binding energy MeV	Observed mass MeV	% difference
proton (P)	$u_bu_bd_b$	985.679	938.261	938.272	−0.0006
neutron (N)	$u_bd_bd_b$	985.508	939.425	939.565	−0.01
Lambda (Λ^0)	$u_bd_bs_b - g^*$	1126.1	1117.7	1115.7	0.18
Sigma (Σ^0)	$u_bd_bs_b$	1196.1		1192.6	0.29
charmed Lambda (Λ_c^+)	$u_bd_bc_b + g^*$	2332.4	2286.3	2286.5	0.005
charmed Sigma (Σ_c^+)	$u_bd_bc_b + 3g^*$	2472.5	2449.4	2452.9	−0.14
bottom Lambda (Λ_b^+)	$u_bd_bb_b$	5631.7	5608.7	5619.4	−0.19
Sigma (Σ^+)	$u_bu_bs_b$	1196.3	1187.9	1189.4	−0.13
Sigma ($\Sigma-$)	$d_bd_bs_b$	1195.9		1197.4	−0.13
charmed Sigma (Σ_c^{++})	$u_bu_bc_b + 3g^*$	2472.7	2449.6	2453.7	−0.18
charmed Sigma (Σ_c^0)	$d_bd_bc_b + 3g^*$	2472.3	2449.4	2453.7	−0.14
bottom Sigma (Σ_b^+)	$u_bu_bb_b + 3g^*$	5842.0	5818.9	5811.3	0.13
bottom Sigma (Σ_b^-)	$d_bd_bb_b + 3g^*$	5841.6	5818.6	5815.5	0.05
Xi (Ξ^0)	$u_bs_bs_b - g^*$	1336.7	1313.6	1319.9	0.38
Xi (Ξ^-)	$d_bs_bs_b - g^*$	1336.5	1328.1	1319.7	−0.15
charmed Xi (Ξ_c^+)	$u_bs_bc_b$	2473.0	2464.6	2467.8	−0.13
charmed Xi (Ξ_c^+)	$d_bs_bc_b$	2472.8		2470.9	0.08
charmed Xi prime ($\Xi_c^{+\prime}$)	$u_bs_bc_b + 2g^*$	2612.9	2581.6	2575.6	0.23
charmed Xi prime ($\Xi_c^{+\prime}$)	$d_bs_bc_b + 2g^*$	2612.9	2581.4	2577.9	0.14
double charmed Xi ($\Xi_c^{++\prime}$)	$u_bc_bc_b + 2g^*$	3679.3	3633.3	3621.4	0.33
bottom Xi (Ξ_b^+)	$u_bs_bb_b$	5842.3	5810.9	5787.8	0.40
bottom Xi (Ξ_b^+)	$d_bs_bb_b$	5842.2	5810.7	5791.1	0.34
charmed Omega (Ω_c^0)	$s_bs_bc_b$	2683.4		2695.2	−0.44
bottom Omega (Ω_b^0)	$s_bs_bb_b$	6052.8		6071	−0.30

The secondary binding energy E_{Q2} for d or u quark is as follows.

$$E_{Q2} = 9E_{Q1}\alpha_w/2 = 1.641\ \text{MeV} \tag{17}$$

The tertiary binding energy E_{Q3} for d or u quark bond is as follows.

$$E_{Q3\ \text{for quark}} = 9E_{Q2}\alpha_w/2 = 0.205\ \text{MeV} \tag{18}$$

The binding energy E_{QQ} for each dd, uu, and du bond is $2E_Q$.

$$E_{QQ} = 2E_Q \tag{19}$$

The mass of neutron (ddu) involves the mass of 2d and u subtracting the binding energy of E_{QQ1} and E_{QQ2} for two quark bonds (2du's) as follows.

$$M_N = Mu + 2Md - 2E_{QQ_{d/u}1} + 2E_{QQ_{d/u}2} = 939.425 \text{ MeV} \tag{20}$$

The calculated mass of neutron is in excellent agreement with the observed value 939.565 MeV with the % mass difference between the calculated and the observed masses = −0.01%.

Proton (duu) is more stable than neutron, so it involves the additional binding energy from the tertiary binding energy E_{QQ3}. For the mass of proton, the baryon number conservation involves the loss of the mass of positron to prevent the decay into positron. Proton becomes permanently stable. The proton mass formula is as follows.

$$M_P = 2Mu + Md - 2E_{QQ_{d/u}1} + 2E_{QQ_{d/u}2} - 2E_{QQ_{d/u}3} - Me = 938.261 \text{ MeV} \tag{21}$$

The calculated mass of proton is in excellent agreement with the observed value 938.272 MeV with the % mass difference between the calculated and the observed masses = −0.0006%.

Being less stable than du bond, the primary binding energy for us bond is one-third of the primary binding energy for du as follows.

$$E_{Q1} = 3M_{\mu_7}\alpha_w / 2 = 4.387 \text{ MeV} \tag{22}$$

The secondary binding energy E_{Q2} for u and s is as follows.

$$E_{Q2} = 3E_{Q1}\alpha_w / 2 = 0.182 \text{ MeV} \tag{23}$$

Only one bond (with binding energy) or less per baryon is allowed for the baryons with s, c, and b. The mass of Sigma (Σ^+) as uus is as follows.

$$M_N = 2Mu + 2Ms - E_{QQ_{us}1} + E_{QQ_{us}2} = 1187.9 \text{ MeV} \tag{24}$$

which is in excellent agreement with the observed value 1189.4 MeV. The binding energy of ds is zero. For example, Sigma (Σ−) with dds has the mass of d + s + d which is 1195.9 MeV in excellent agreement with the observed 1197.4 MeV.

In the two baryons with the same quark composition, the difference in the masses between the two baryons is equal to the multiple of g*, and one baryon has morebond (with binding energy) than the other baryon, so a bond is added or subtracted in one of the two baryons. For example, the two baryons, Lambda (Λ^0) and Sigma (Σ^0), are uds. Lambda (Λ^0) has the mass of u + d + s − g* − E_{QQus1} + E_{QQus2} which is 1117.7 MeV in excellent agreement with the observed 1115.7 MeV. One bond is subtracted in Sigma (Σ^0) which has the mass of u + d + s which is 1196.1 MeV in excellent agreement with 1192.6 MeV.

The binding energies of dd, uu, du, uc, and ub are the same. The binding energies of ds, dc, and db are zero. For example, bottom Lambda (Λ_b^+) with udb has the mass of u + d + b − E_{QQub1} + E_{QQub2} which is 5608.7 MeV in excellent agreement with the observed 5619.4 MeV. Only one bond or less is allowed for the baryons with s, c, and b except in the two baryons with the same quark composition where one bond is added in one of the two baryons. For example, the two baryons, charmed Sigma (Σ_c^+) and charmed Lambda (Λ_c^+), are udc, Charmed Sigma (Σ_c^+) has the mass of u + d + c + 3g* − E_{QQuc1} + E_{QQuc2} which is 2449.4 MeV in excellent agreement with the observed 2452.9 MeV. One bond is added in charmed Lambda (Λ_c^+) which has the mass of u + d + c + g* − E_{QQuc1} + E_{QQuc2} − E_{QQud1} + E_{QQud2} which is 2286.3 MeV in excellent agreement with the observed 2286.5 MeV.

Without d/u, the baryons have no binding energy. For example, charmed Omega (Ω_c^0) with ssc has the mass of s + s + c which is 2683.4 MeV which is in excellent agreement with the observed 2695.2 MeV.

The calculated masses and the observed masses [10] of mesons are in **Table 5**. Since parallel spins have higher energy than antiparallel spins, vector mesons with parallel spins have higher masses than pseudoscalar mesons with antiparallel spins. For higher-mass vector mesons (parallel spins), vector t, b, c, and s are baryonic t, b, c, s, and vector d and u are baryonic d and u plus basic gluon (g*). For lower-mass pseudoscalar mesons (antiparallel spins), pseudoscalar t is a part of vector t, pseudoscalar b and c are vector quarks minus g*, and pseudoscalar d, u, and s are derived from g*.

The mass of $\pi^\pm$ is the mass of 2g* minus the mass of $e^\pm$ as proposed by Peter Cameron [17]. The calculated mass of $\pi^\pm$ is 139.5395 MeV which is in excellent agreement with the observed 139.5702 MeV. $\pi^\pm$ has much longer mean lifetime than other mesons to indicate that the composite of $\pi^\pm$ is not normal composite of u and d quarks. Another pseudoscalar meson with long mean lifetime is K^+ ($u\bar{s}$) which has the composition of 7g* + 7e with the calculated mass of 493.754 MeV in excellent agreement with the observed 493.677 MeV.

The mass of π^0 involves the composite of pseudoscalar u and pseudoscalar d quarks as $(u_p\bar{u}_p + d_p\bar{d}_p)/2$. The binding energy for pseudoscalar meson involves the auxiliary mass dimensional orbital at d = 7 similar to the binding energy in u and d quarks for baryons as Equation (16).The binding energy for pseudoscalar u and d does not involve 3 colors as in 3-color gluons, so similar to Equation (16), the primary binding energy for pseudoscalar u and d quarks at d = 7 with α_w as follows.

$$E_{Q1} = 3M_{g*}\alpha_w/2 = 2.911\,\text{MeV} \tag{25}$$

The secondary binding energy is as follows.

$$E_{Q2} = 3E_{Q1}\alpha_w/2 = 0.121\,\text{MeV} \tag{26}$$

The binding energy E_{QQ} for each dd, uu, and du bond is $2E_Q$. π^0 is $(u_p\bar{u}_p + d_p\bar{d}_p)/2$, so similar to Equation (20), the mass of π^0 is as follows.

$$M_{\pi^0} = Mu_p + Md_p - E_{QQ1} + E_{QQ2} = 134.982\text{ MeV} \tag{27}$$

which is in excellent agreement with the observed value 134.9766 MeV.

The binding energy for the ds bond and the us bond is three times of the d/u quark bond to form the composite boson with three flux quanta as follows.

$$E_{Q1} = 3\times 3M_{g*}\alpha_w/2 = 8.732\text{ MeV} \tag{28}$$

The secondary binding energy is as follows.

Table 5. The masses of mesons and the Higgs bosons.

Meson	Spin	Composition	Calculated mass without binding energy (MeV)	Calculated mass with binding energy (MeV)	Observed mass (MeV)	% difference
pion ($\pi^{\pm}$)	0	$2g^* - e$ as $u\bar{d}$	139.540		139.570	−0.02
pion (π^0)	0	$(u_p\bar{u}_p + d_p\bar{d}_p)/2$	140.562	134.982	134.977	0.004
charged rho meson (ρ^+)	1	$u_v\bar{d}_v$	797.11	774.07	775.11	−0.13
omegameson (ω)	1	$u_v\bar{d}_v$	797.11	785.59	782.65	0.32
eta meson (η)	0	$(u_p\bar{u}_p + d_p\bar{d}_p + s_p\bar{s}_p)/2$	563.949	548.663	547.862	0.15
eta prime meson (η')	0	$(u_p\bar{u}_p + d_p\bar{d}_p)/2 + s_p\bar{s}_p$	987.34	956.76	957.78	−0.11
kaon (K^+)	0	$7(g^* + e)$ as $u_p\bar{s}_p$	493.754		493.677	0.02
kaon (K^0)	0	$7(g^* + 2e)$ as $d_p\bar{s}_p$	497.331		497.614	−0.06
kaon (K^{*+})	1	$u_v\bar{s}_v$	937.7	891.60	891.66	−0.007
phi meson (ϕ)	1	$s_v\bar{s}_v$	1078.1	1018.9	1019.5	−0.06
D meson (D^{*0})	1	$c_v\bar{u}_v$	2004.0		2007.0	−0.15
strange D meson (D_s^+)	0 0	$e + 20(3g^*/2) - 2g^*$ $(c_p\bar{s}_p)$	1961.2 (1958.7)		1968.3	−0.36
strange D meson (D_s^{*+})	1 1	$e + 20(3g^*/2)$ $(c_v\bar{s}_v)$	2101.3 (2144.4)		2112.1	−0.52
charmed eta meson (η_c)	0 0	$2e + 30(3g^*/2) - 2g^*$ $(c_p\bar{c}_p)$	3012.1 (3070.6)		2983.6	0.95
J/Psi	1 1	$2e + 30\ (3g^*/2) - g^*$ $(c_v\bar{c}_v)$	3082.1 (3210.6)		3096.9	−0.48
D meson (D^0)	0	$c_p\bar{u}_v - g^*$	1863.9		1864.8	−0.05
D meson (D^{*0})	1	$c_v\bar{u}_v$	2003.8		2010.3	−0.32
B meson (B_d^0)	0	$d_v\bar{b}_p$	5303.1		5279.6	0.44
B meson (B_d^{*0})	1	$d_v\bar{b}_v$	5373.1		5325.2	0.89
B meson (B_c)	0 0	$e + 60(3g^*/2)$ $(c_p\bar{b}_p)$	6302.8 (6440.0)		6275.6	0.41
bottom eta meson (η_b)	0 0	$2e + 90(3g^*/2) - g^*$ $(b_p\bar{b}_p)$	9384.4 (9809.3)		9398.0	−0.14
upsilon meson (γ)	1 1	$2e + 90(3g^*/2)$ $(b_v\bar{b}_v)$	9454.4 (9949.4)		9460.3	−0.06
top-bottom quark-antiquark composite	0	$b_p\bar{b}_p + (b_p\bar{b}_p + t_p\bar{t}_p)/2$	27.9 GeV		28 GeV	−0.3
pseudoscalartop quark-antiquark (absent)	0	$t_p\bar{t}_p$	26.4 GeV		not observed	
t quark	½	$t_7 + t_8$	175.4 GeV		172.4 GeV	1.71
perturbative Higgs boson (absent)	0	$W^+W^-Z^0$	252 GeV		not observed	
low Higgs boson	0	$W^+W^-Z^0/2$	126 GeV		125 GeV	0.79
high Higgs boson	0	$3W^+W^-Z^0$	756 GeV		750 GeV	0.8

$$E_{Q2} = 9E_{Q1}\alpha_w/2 = 1.089\ \text{MeV} \tag{29}$$

The mass of η as ($u_p\bar{u}_p$ + $d_p\bar{d}_p$ + $s_p\bar{s}_p$)/2 is as follows.

$$M_\eta = Mu_p + Md_p + Ms_p - E_{QQ1} + E_{QQ2} = 548.663\ \text{MeV} \tag{30}$$

which is in excellent agreement with the observed value 547.862 MeV.

The mass of η' as ($u_p\bar{u}_p$ + $d_p\bar{d}_p$)/2 + $s_p\bar{s}_p$ is as follows.

$$M_{\eta'} = Mu_p + Md_p + 2Ms_p - 2E_{QQ1} + 2E_{QQ2} = 956.764\ \text{MeV} \tag{31}$$

which is in excellent agreement with the observed value 957.78 MeV.

The mesons with c, b, and t have no binding energy. For d_p and u_p in the pseudoscalar mesons with c and b, d_p = d_b = d_v − g*, and u_p = u_b = u_v − g*. For example, the mass of D meson (D^0) with c_p $\bar{u}_p$ is c_p + u_p = c_p + u_v − g* which is 1863.9 MeV in excellent agreement with the observed 1864.8 MeV.

The binding energy for vector dd, uu, and du bonds involves the same binding energy as baryonic d/u quark bond as Equations (16) and (17), so vector ρ^+ as $u_v\bar{d}_v$ with the binding energy derived from Equations (16) and (17) is as follows.

$$M_\rho = Mu_v + Md_v - E_{QQ1} + E_{QQ2} = 774.07\ \text{MeV} \tag{32}$$

which is in excellent agreement with the observed value 775.11 MeV. As in the baryons with the same quark composition, charged rho meson (ρ^+) and omega-meson (ω) have the same composition as $u_v\bar{d}_v$, so 1/2 bond is subtracted in omegameson (ω) which has the mass as follows.

$$M_\omega = Mu_v + Md_v - E_{Q1} + E_{Q2} = 785.59\ \text{MeV} \tag{33}$$

which is in excellent agreement with the observed value 782.65 MeV.

The binding energy for vector ds bond and us bond is twice of the binding energy for d/u quarks. The mass for kaon (K^{*+}) with $u_v\bar{s}_v$ is as follows.

$$\begin{aligned} M_{K^{*+}} &= Mu_v + Ms_v - 2E_{QQ1} + 2E_{QQ2} \\ &= Mu_v + Ms_v - 52.648\ \text{MeV} + 6.565\ \text{MeV} \\ &= 891.60\ \text{MeV} \end{aligned} \tag{34}$$

which is in excellent agreement with the observed value 891.66 MeV. The binding energy for vector ss bond has the opposite sign for E_{QQ2}, so phi meson (ϕ) as $s_v\bar{s}_v$ has the mass of s_v + s_v—52.648 MeV - 6.565 MeV which is 1018.9 MeV in excellent agreement with the observed 1019.5 MeV. The mesons with c, b, and t have no binding energy. For example, D meson (D^{*0}) with $c_v\bar{u}_v$ has the mass of c_v + u_v which is 2004.0 MeV which is in excellent agreement with the observed 2007.0 MeV.

The masses of the mesons of c/b without d/u follow the meson mass formula by Malcolm H. MacGregor [18] to match the masses of mesons derived from the quark mass formula as Equation (15). The MacGregor's meson mass formula derived from the muon mass formula as Equation (4) is as follows.

$$M_{meson} = Me \text{ or } 2Me + 2n\left(3M_{g^*}/2\right) \tag{35}$$

where one e is for charge meson and 2e for neutral meson, and n (integer) is determined by the masses of mesons calculated from the quark mass formula as

Equation (15). For example, the calculated mass of vector strange D meson (D_s^{*+}) as $c_v\bar{s}_v$ is $c_v + s_v$ which is 2144.4 MeV. To match 2144.4 MeV, the MacGregor's meson mass formula generates e + 20(3g*/2) which is 2101.3 MeV in excellent agreement with the observed 2112.1 MeV. The mass different between pseudoscalar strange D meson (D_s^{+}) as $c_p\bar{s}_p$ and vector strange D meson (D_s^{*+}) as $c_v\bar{s}_v$ is 2g*. Strange D meson (D_s^{+}) has the mass of e + 20(3g*/2) – 2g* which is 1961.2 MeV in excellent agreement with the observed 1968.3 MeV.

The MacGregor's meson mass formula in Equation (35) for the mesons of c/b without d/u/s, n = the multiple of 3to simulate baryonic quark which uses 3μ as in Equation (9). For example, vector upsilon meson (γ) with $b_v\bar{b}_v$ has the mass of 2e + 90(3g*/2) which is equal to 9454.4 MeV in excellent agreement with the observed 9460.3 MeV. The mass difference between vector $b\bar{b}$ and pseudoscalar $b\bar{b}$ is the mass of g*. Pseudoscalar bottom eta meson (η_b) with $b_p\bar{b}_p$ has the mass of 2e + 90(3g*/2) – g* which is equal to 9384.4 MeV in excellent agreement with the observed 9398.0 MeV.

Pseudoscalar and partial t quark is t_7, while vector and full t quark is $t_7 + t_8$. Vector and full t quark with enormous mass is extremely short-lived, so top quark-antiquark does not have time before they decay to form hadrons, resulting in "bare" t quark and antiquark. The calculated mass of t is 175.4 GeV in good agreement with the observed 172.4 GeV.

The summary of binding energies in hadrons is in Table 6. The binding energies are derived from the auxiliary mass dimensional orbital at d = 7 as in Equations (16), (17), (18). (19), (22), (23), (25), (26), (28), (29), (32), and (34). In general, the relatively stable hadrons with d, u, and s quarks have binding energies, while relatively unstable hadrons with c, b, and t quarks and without u and d do not have binding energies. The baryons with the u and s/c/b bonds have binding energies, and the baryons with the d and s/c/b bonds do not have binding

Table 6. The binding energies in hadrons.

	Bond (QQ)	Primary (MeV)	Secondary (MeV)	Tertiary (MeV)	Equation
Baryon	dd, uu, du, uc, ub	26.324	3.282	0.409	16, 17, 18. 19
	us	8.775	0.365	0	22, 23
	ds, dc, db, baryons without d and u	0	0	0	
Pseudoscalar meson	dd, uu, du	5.822	0.242	0	25, 26
	ds, us	17.465	2.178	0	28, 29
	mesonswith c, b, and t	0	0	0	
Vector meson	dd, uu, du	26.324	3.282	0	32
	ds, us, ss	52.648	6.565	0	34
	mesons with c, b, and t	0	0	0	

energies. The mesons with c, b, and t quarks do not have binding energies. Pseudoscalar mesons have lower binding energies than baryons and vector mesons. In the two hadrons with the same quark composition, one hadron has more bond (with binding energy) than the other, so a bond is added or subtracted in one of the two hadrons.

4. The Top-Bottom Quark-Antiquark Composite

In the search for resonances produced in association with a b quark jet and a second jet, and decaying to a muon pair, the CMS Collaboration at the LHC recently reported an excess of events above the background near a dimuon mass of 28 GeV. The search is carried out in two categories from proton-proton collisions at center-of-mass energies of 8 and 13 TeV. The first category involves a b quark jet in the central region and at least one jet in the forward region, while the second category involves two jets in the central region, at least one of which is identified as a b quark jet, no jets in the forward region. At the 8 TeV collision, the first category has 4.2 standard deviation, while the second category has 2.9 standard deviations. At the 13 TeV collision, the first category has 2.0 standard deviations, while the second category results in a 1.4 standard deviation deficit.

As shown in Figure 1, Table 1, Table 3, and Table 5, pseudoscalar t quark is t_7 (13.2 GeV), while vector and full t quark (175.4 GeV) is $t_7 + t_8$. This paper posits that the resonance with 28 GeV observed recently at the LHC is the pseudoscalar top-bottom quark-antiquark composite which has the calculated mass of 27.9 GeV derived from the periodic table of elementary particles in good agreement with the observed 28 GeV as shown in Table 5. The calculated mass is the mass of three pseudoscalar b quarks and one pseudoscalar t quark which represent the composite of $b_p\bar{b}_p + (b_p\bar{b}_p + t_p\bar{t}_p)/2$. As described in the periodic table of elementary particles, pseudoscalar t quark is only a part of full t quark, so pseudoscalar $t_p\bar{t}_p$ (26.4 GeV) cannot exist independently, and can locate within a composite, such as the top-bottom quark-antiquark composite. As shown in the observation at the LHC, the resonance with 28 GeV weakens significantly at the higher energy collision (13 TeV), because at the higher collision energy, low-mass pseudoscalar $t\bar{t}$ in the composite likely becomes independent full high-mass vector $t\bar{t}$ moving out of the composite. The presence of the top-bottom quark-antiquark composite weakens, disappears, or gets inverted at 13 TeV as shown at the LHC. Normally, the presence of resonance gets stronger at the higher collision energy.

To account for the observed two jets, the composite has two jets consisting of a $b\bar{b}$ jet and a b + t jet for $(b_p\bar{b}_p + t_p\bar{t}_p)/2$, where $b\bar{b}$ jet is more stable than b + t jet which decays faster into the jet in the forward region to constitute the first category of the search by the CMS Collaboration at the LHC. Since $t_p\bar{t}_p$ is less stable than $b_p\bar{b}_p$, so the decay of the b + t jet is faster to allow the greater standard deviations for the first category than for the second category. The sum of the standard deviations from both categories is greater than 5.

5. The Higgs Boson Doublet

One important open theoretical issue about the Higgs boson is the triviality problem [19]. Within the perturbation theory, the Higgs boson mass squared is proportional to the self-coupling. However, the scalar self-coupling for the scalar Higgs boson leads to triviality or non-interaction which is inconsistent to the interactive Higgs boson. To deal with the triviality problem, Cea and Cosmai [20] [21] established the non-perturbation non-trivial rescaling of the Higgs condensate to avoid the vanishing self-coupling, resulting in the generation of the heavy Higgs boson with 754 GeV. According to Cea, the theoretical expectations of the predicted heavy Higgs boson (754 GeV) are in fairly good agreement with the observations at the LHC Run 2with an estimated statistical significance of more than five standard deviations [22].

This paper proposes that derived from the non-trivial rescaling of the Higgs condensate, the Higgs boson doublet consists of the high Higgs boson from the upward rescaling of the Higgs condensate and the low Higgs boson from the downward rescaling. The perturbative Higgs boson became the non-perturbative Higgs boson doublet irreversibly during the spontaneous symmetry breaking. The observed mass of the high Higgs boson is 750 GeV [22] [23] [24] [25], and the observed mass of the low Higgs boson is 125 GeV [26] [27].

This paper also proposes that the Higgs bosons are the intermediate vector boson composites whose condensate provides the masses directly to the intermediate vector bosons during the spontaneous symmetry breaking. The Higgs bosons consist of the perturbative Higgs boson ($W^+W^-Z^0$ = 252 GeV = M_{W+} + M_{W-} + M_{Z0}), the low Higgs boson ($W^+W^-Z^0/2$ = 126 GeV) from the downward rescaling, and the high Higgs boson ($3W^+W^-Z^0$ = 756 GeV) from the upward rescaling. The perturbative Higgs boson is absent, and there is no Higgs boson pair [28]. The low Higgs boson as $W^+W^-Z^0/2$ with respect to $W^+W^-Z^0$ is like one quark in a di-quark meson, while the high Higgs boson as $3W^+W^-Z^0$ with respect to $W^+W^-Z^0$ is like one baryon consisting of three quarks. In this way, the Higgs boson doublet from the rescaling provides the structures of the quark compositions for mesons and baryons which are not like leptons without composite structures. The calculated masses (126 GeV and 756 GeV) of the Higgs boson doublet are in excellent agreements with the observed masses (125 GeV and 750 GeV) [22]-[27]. For the periodic table of elementary particles, the Higgs mechanism assigns the mass of B_7 as the mass of Z^0. The mass of B_7 produces α_w which determines the masses of quarks as in the quark mass formula in Equation (15).

6. Summary

This paper posits that the observed resonance with 28 GeV at the LHC is the pseudoscalar top-bottom quark-antiquark composite which has the calculated mass of 27.9 GeV derived from the periodic table of elementary particles. The calculated mass is the mass of three pseudoscalar b quarks and one pseudoscalar t quark to represent the composite of $b_p\overline{b}_p$ + ($b_p\overline{b}_p$ + $t_p\overline{t}_p$)/2 where p = pseudoscalar. (The quark in pseudoscalar meson is denoted as "pseudoscalar quark", while the quark in vector mesons is denoted as "vector quark" which has higher mass than pseudoscalar quark.) In the periodic table of elementary particles, pseudoscalar t quark (13.2 GeV) is only a part of full t quark (175.4 GeV), so pseudoscalar $t_p\overline{t}_p$ (26.4 GeV) cannot exist independently, and can exist only in the top-bottom quark-antiquark composite. As shown in the observation at the LHC, the resonance with 28 GeV weakens significantly at the higher energy collision (13 TeV), because at the higher collision energy, low-mass pseudoscalar $t_p\overline{t}_p$ in the composite likely becomes independent full high-mass vector $t_v\overline{t}_v$ moving out of the composite. To account for the observed two jets, the composite has two jets consisting of a $b\overline{b}$ jet and a b + t jet, where $b\overline{b}$ jet for ($b_p\overline{b}_p$ + $t_p\overline{t}_p$)/2 is more stable than b + t jet which decays faster into the jet in the forward region to constitute the first category of the search by the CMS Collaboration at the LHC.

The periodic table of elementary particles is based on the seven mass dimensional orbitals derived from the seven extra dimensions of 11 spacetime dimensional membrane particles. The seven mass dimensional orbitals include the seven principal mass dimensional orbitals for stable baryonic matter leptons (electron and neutrinos), gauge bosons, gravity, and dark matter and the seven auxiliary mass dimensional orbitals for unstable leptons (muon and tau) and quarks, and calculate accurately the masses of all elementary particles and the cosmic rays by using only five known constants. For baryons, the quarks in the periodic table of elementary particles are baryonic quarks. For high-mass vector mesons (parallel spins), vector t, b, c, and s are baryonic t, b, c, s, and vector d and u are baryonic d and u plus basic gluon (g^* = 70 MeV) which has the mass of electron/α where α is the fine structure constant of electromagnetism. For low-mass pseudoscalar mesons (antiparallel spins), pseudoscalar t is a part of vector t, pseudoscalar b and c are vector quarks minus g^*, and pseudoscalar d, u, and s are derived from g^*. The binding energies among quarks are derived from the auxiliary mass dimensional orbital. With these masses and binding energies of quarks, the masses of hadrons can be calculated in excellent agreement with the observed masses of hadrons by using only five known constants in the periodic table of elementary particles. For examples, the calculated masses of proton, neutron, pion ($\pi^{\pm}$), and pion (π^0) are 938.261, 939.425, 139.540, and 134.982 MeV in excellent agreement with the observed 938.272, 939.565, 139.570, and 134.977 MeV, respectively with 0.0006%, 0.01%, 0.02%, and 0.004%, respectively for the difference between the calculated and observed mass. The calculated masses of the Higgs bosons as the intermediate vector boson composites are in excellent agreements with the observed masses. In conclusion, the calculated masses of the top-bottom quark-antiquark composite (27.9 GeV), hadrons, and the Higgs bosons by the periodic table of elementary particles are in excellent agreement with the observed masses of resonance with 28 GeV at the LHC, hadrons, and the Higgs bosons, respectively.

Conflicts of Interest

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

References

[1] CMS Collaboration (2018) Search for Resonances in the Mass Spectrum of Muon Pairs Produced in Association with b Quark Jets in Proton-Proton Collisions at s √ = 8 and 13 TeV. arXiv:1808.01890 [hep-ex].

[2] Chung, D. (1998) *Speculations in Science and Technology*, **21**, 277-289. https://doi.org/10.1023/A:1005513404873

[3] Chung, D. (2014) *Journal of Modern Physics*, **5**, 1234-1243. https://doi.org/10.4236/jmp.2014.514123

[4] Chung, D. (2018) *Journal of Modern Physics*, **9**, 2308-2319. https://doi.org/10.4236/jmp.2018.913146

[5] Chung, D. (2016) *Journal of Modern Physics*, **7**, 1591-1606. https://doi.org/10.4236/jmp.2016.712144

[6] Chung, D. (2014) *Journal of Modern Physics*, **5**, 1467-1472. https://doi.org/10.4236/jmp.2014.515148

[7] Fujita, J. and Miyazawa, H. (1957) *Progress of Theoretical Physics*, **17**, 360-365. https://doi.org/10.1143/PTP.17.360

[8] Kamilla, R., Wu, X. and Jain, J. (1996) *Physics Review Letters*, **76**, 1332. https://doi.org/10.1103/PhysRevLett.76.1332

[9] Jain, J. (2007) Composite Fermions. Cambridge University Press, New York. https://doi.org/10.1017/CBO9780511607561

[10] Particle Data Group (2018) *Physical Review D*, **98***,* 030001. https://doi.org/10.1103/PhysRevD.98.030001

[11] Barut, A. (1979) *Physical Review Letters*, **42**, 1251. https://doi.org/10.1103/PhysRevLett.42.1251

[12] Krasznahorkay, A. et al. (2016) *Physical Review Letters*, **116**, Article ID: 042501. https://doi.org/10.1103/PhysRevLett.116.042501

[13] Feng, J. et al (2016) *Physical Review Letters*, **117**, Article ID: 071803. https://doi.org/10.1103/PhysRevLett.117.071803

[14] De Rujula, A., Georgi, H. and Glashow, S. (1975) *Physics Review D*, **12**, 147. https://doi.org/10.1103/PhysRevD.12.147

[15] Griffiths, D. (2008) Introduction to Elementary Particles. Wiley-VCH, Weinheim.

[16] El Naschie, M. (2002) *Chaos, Solitons and Fractals*, **14**, 369-376. https://doi.org/10.1016/S0960-0779(02)00022-X

[17] Cameron, P. (2011) *Apeiron*, **18**, 29-42.

[18] MacGregor, M. (2007) The Power of Alpha: The Electron Elementary Particle Generation with Alpha-Quantized Lifetimes And Masses. World Scientific Publishing, Singapore.

[19] Callaway, D.J.E. (1988) *Physics Reports*, **167**, 241-320. https://doi.org/10.1016/0370-1573(88)90008-7

[20] Cea, P. and Cosmai, L. (2012) *International Scholarly Research Notices High Energy Physics*, **2012**, Article ID: 637950.

[21] Cea, P. (2017) The HT Higgs Boson at the LHC Run 2.

[22] Cea, P. (2018) Evidence of the True Higgs Boson HT at the LHC Run 2.

[23] ATLAS and CMS (2015) ATLAS and CMS Physics Results from Run 2. https://indico.cern.ch/event/442432/

[24] The CMS Collaboration (2018) Measurements of Properties of the Higgs Boson in the Four-Lepton Final State at s = 13 TeV, CMS PAS HIG-18-001.

[25] The ATLAS Collaboration (2017) Search for Heavy *ZZ* Resonances in the $\ell+\ell-\ell+\ell-$ and $\ell+\ell-\nu\bar{\nu}$ Final States Using Proton Proton Collisions at s =13 TeV with the ATLAS Detectorar.

[26] The CMS Collaboration (2012) *Physics Letters B*, **716**, 30-61. https://doi.org/10.1016/j.physletb.2012.08.021

[27] The ATLAS Collaboration (2012) *Physics Letters B*, **716**, 1-29.

[28] ATLAS Collaboration (2018) *The Journal of High Energy Physics*, **11**, 040.

Journal of Modern Physics, 2018, 9, 2657-2724
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Water Memory Due to Chains of Nano-Pearls

Auguste Meessen

Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium
Email: auguste@meessen.net

How to cite this paper: Meessen, A. (2018) Water Memory Due to Chains of Nano-Pearls. *Journal of Modern Physics*, **9**, 2657-2724.
https://doi.org/10.4236/jmp.2018.914165

Received: November 20, 2018
Accepted: December 26, 2018
Published: December 29, 2018

Abstract

Biologically active molecules create substitutes in liquid water by forming *single-domain ferroelectric crystallites.* These nanoparticles are spherical and constitute growing chains. The dipoles are aligned, but can be set in oscillation at the frequency of vibration of the charged part of active molecules. They are then automatically trimmed and become information carriers. Moreover, they produce an oscillating electric field, causing autocatalytic multiplication of identical chains in the course of successive dilutions. Active molecules are thus only required to initiate this process. Normally, they excite their specific receptors by resonance, but trimmed chains have the same effect. This theory is confirmed by many measurements.

Keywords

Water Memory, Extra-High Dilutions, Ferroelectric Particles, Chains of Nanoparticles, Water Bridges, Molecular Interactions

1. Introduction

The concept of water memory is based on experimental results of measurements, published 30 years ago in the prestigious scientific journal Nature [1]. This astonishing phenomenon had been discovered by *Jacques Benveniste* and his collaborators, but one month later, the same journal declared that it was a delusion [2]. Even the publication of the discovery was already accompanied by an astounding editorial reservation [3]. The editor in chief, *John Maddox*, declared indeed that "there is no physical basis for such an activity" and announced even that independent investigators would "observe repetitions of the experiments".

When they arrived at Benveniste's laboratory, it turned out that John Maddox was accompanied by the professional magician *James Randi* and the debunker of scientific fraud *Walter Steward.* The objective was thus to detect errors or fraud. During their stay in Paris, the first experiments confirmed the published result, but not later ones. The inquisitors did immediately publish a devastating conclusion: the reported results are not reproducible and merely due to imagination [2]. Randi stated even that they should be compared to the sensational claim of having seen a unicorn, where there was merely a goat.

The load of the accusation fell on Jacques Benveniste (1935-2004). He was a medical doctor, who had practiced several years before opting for research. In California, he discovered the platelet-activating factor and determined its role in immunology [4] [5]. He returned to France in 1973 and joined INSERM (National Institute of Health and Medical Research). Since 1980, he directed Unit 200, devoted to research in immunology, allergy and inflammation. He discovered there a very sensitive method to detect allergens. It required careful counting of discolored cells, but *Elisabeth Davanas* succeeded very well [1]. The young medical doctor *Bernard Poitevain,* who joined the team to prepare a thesis, asked somewhat later if he could test the efficiency of homeopathic dilutions by means of this method. Benveniste answered: "try if you want, but there will be no effect; high dilutions are merely water" ([6], p. 45). Nevertheless, the results were positive and confirmed by other members of the group.

This phenomenon was totally unexpected, since successive dilutions of biologically active molecules in pure water do necessarily lead to their complete elimination. Was it really possible to create substitutes that are only constituted of water molecules? Further experimentation did prove that these structures should even be able to mimic active molecules of different types. No one knew how this might be achieved. Benveniste verified therefore if these hypothetical structures could be destroyed. It appeared that after heating extra-high dilutions (EHDs) of histamine during one hour at 70˚C, they had completely lost their biological efficiency. Exposition to ultrasound had the same effect ([6], pp. 53-54). The puzzling phenomenon of "water memory" had even a characteristic property: the biological efficiency of EHDs decreased at first, but increased again and reached a high value after about the 9th decimal dilution. Then it dropped and continued to vary in a quasi-periodic way during successive dilutions.

Since the reality of this phenomenon was tested many times for different substances, Benveniste thought that these facts had to be published, although they were unexplained. He insisted on the observed quasi-periodic variations, by providing two figures [1]. However, Maddox was convinced that the reported results cannot be real. He required independent confirmations. They were provided by other laboratories in Italy, Canada and Israel. Eventually, after two-year long discussions ([6], p. 51), Maddox accepted to publish the article if Benveniste did agree that a team of experts could come to "verify the quality of the experimentation". This was more than bizarre, but Benveniste had nothing to hide. The outcome was the terrible accusation that we mentioned [2]. Benveniste was allowed to add a reply, but could only describe how the investigators had proceeded. They created a climate of "intense and constant suspicion". Actually, Benveniste and his collaborators were treated like "criminals".

In their report, the group of inquisitors insisted on the variability of the peaks of activity and claimed that measurements had been treated in "disregard of statistical principles". They declared even that "the laboratory has fostered and then cherished a delusion about the interpretation of this data". *Melinda Baldwin,* lecturer on History of Science at Harvard University, identified the actual cause of this grave incident. Maddox considered that scientific journals are shaping science by controlling its quality [7]. This implies enormous power, but also the danger to defend orthodoxy by condemning any deviation. Baldwin mentioned the quasi-periodic variations of the reported biological efficiency. They were strange, but could have an intrinsic cause, deserving further research.

Benveniste considered that his duty was only to establish the reality of these "unbelievable and fear-provoking" facts, since their meaning and the underlying mechanism could be studied later on ([6], p. 61). This is not at all unusual in science. After the events of 1988, the experiments were repeated many times with the statistician *Spira*, but the previous results were validated. In the course of further tests, Benveniste made a second discovery. He realized in 1993 that two EHDs were able to perturb one another and thought that this was due to electromagnetic fields. He suspected that these "signals" could play an essential role in the constitution of water memory and did prove that they can be detected, amplified, recorded and transmitted. They seemed to be noise in the frequency domain of audible sound waves (lower that 20 kHz). However, when pure water was exposed to them, it acquired the same properties as EHDs of the initially dissolved active molecules. This *information transfer* confirmed that water memory is real, but INSERM did not renew the contract for Unit 200 in 1995. Benveniste pursued his research in a room on the parking of his former institute, with a caravan for storing materials. He focused now his efforts on developing "digital biology".

Alain Kaufmann presented in 1994 basic facts and an analysis of the sociological context [8]. *Michel Schiff* published the same year a more detailed description of the experiments and denounced the dangers of censorship [9]. Benveniste, who had reluctantly accepted to write the preface of this book, attributed there the radical rejection of available evidence to the following reasons: "It *cannot be,* since if it were true it would have been found two hundred years ago" and "there is *no theory* behind it." We can add that the discovery of an anomaly may lead at first to incredulity, but well-documented facts should induce a search of their cause. It might be necessary, indeed, to correct some previous assumptions. Peer evaluation is necessary and useful, but not infallible. Schiff insisted that even scientists can "suppress unwanted knowledge" when it would "shatter their current beliefs". Actually, "the long history of scientific dogmatism shows that today's heresy could well become tomorrow's scientific truth".

Francis Beauvais, a former collaborator of Benveniste, provided in 2007 much more details on events at that time [10]. *Yolène Thomas* [11] continued research on water memory, but it had to be camouflaged as concerning properties of EHDs. Even the French virologist *Luc Montagnier,* who received in 1988 the

Nobel Prize for his contribution to the discovery of the HIV virus, was violently attacked when he resumed the experimental study of water memory. He was motivated by scientific curiosity and the constant need of improved medical treatments, while his opponents negated *a priori* that EHDs in pure liquid water could modify this solvent. Montagnier presented his experimental method in a documentary, realized in 2014 by French TV [12]. It is also available in English [13] and does clearly demonstrate that water memory involves detectable signals at frequencies like those of audible sound waves (20 - 20,000 Hz). Montagnier published in 2009 and 2010 important results concerning water memory [14] [15]. He discovered even that viral DNA sequences can be reproduced by means of transmitted signals when the building blocks are available in pure water. These signals had thus to provide the required master plan and this fact might account for strange resurgences of some sicknesses. The conference of Montagnier at UNESCO in 2014 stressed this fact and other medical applications [16].

Visceral opposition to the concept of water memory was often motivated by fear that it could justify homeopathy. The aim of the present study is merely to find out if water memory is real or not. This has to be viewed as a basic problem for condensed matter physics, since bonds between water molecules are constantly broken by thermal agitation in liquid water at the time scale of picoseconds. *Martin Chaplin,* specialist of properties of water molecules, proposed therefore in 2007 that water memory could result from creating statistically stable clusters of water molecules [17]. Individual molecules would there be replaced by other ones without affecting the global structure. This hypothesis was the most plausible one, but Martin Chaplin added that "much research work remains to be undertaken if these real and observable facts are to be completely understood".

The structure of this article results from the itinerary that we followed. In Section 2, we examine the internal structure of water molecules and their possible interactions. This leads to the concept of "water pearls". In Section 3, we explain why biologically active molecules can create chains of these nano-pearls and why they account for water memory. Section 4 presents more observational evidence concerning these chains. It is diverse, detailed and very remarkable, but the concept of water pearls accounts for known facts, while the alternative concept of Coherent Domains does not. In Section 5, we insist on the most important consequence of water memory: molecular interactions are not only possible by means of the "key and slot" model of chemical reactions. Intermolecular communications can also result from oscillating electric fields and resonance effects. It will appear once more that "Nature is written in Lingua Mathematica", as Galilei stated already, but we endeavor to be understandable by non-specialists.

2. Interactions of Water Molecules

2.1. Their Structure and the Dipole Approximation

Martin Chaplin provides detailed information about the internal structure of water molecules [18] and various models that have been proposed [19]. The usual "stick and ball model" insists on the chemical composition (H_2O), but the protons of both light atoms are deeply embedded in the common electron cloud of the oxygen and hydrogen atoms. Water molecules are thus practically spherical, but at close range, the protons are surrounded by a spherically symmetric excess density of electrons. They are thus equivalent to point-like charges $q \approx e/3$. The core of the oxygen atom and the remaining part of the electron cloud are equivalent to a central point-like charge $-2q$.

The kinetic diameter (for collisions) of water molecules in the terrestrial atmosphere is 0.265 nm. H_2O molecules are thus smaller than O_2, N_2, CO_2 and H_2 [20]. Their size is slightly greater in the liquid state, because of interactions with surrounding water molecules. The average separation of two oxygen atoms is there measurable by means of x-ray diffraction. The resulting diameter is d = 0.275 nm. In the gaseous state, the angle $^{H}O^{H}$ is 104.5˚, but it is close to 106˚ in the liquid state. Vibrational and rotational spectra of water molecules disclosed that the length of OH bonds is $\delta \approx 0.095$ nm. Thus, $\delta/d \approx 1/3$. Since water molecules behave in the liquid state like hard spheres that can easily roll on one another, we adopt the model of **Figure 1**. To simplify later calculations, we chose natural units, where $\delta = 1$ and $q = 1$. The diameter of a water molecule in the liquid state is then $d \approx 2.90$. Since the angle $\varphi = 53˚$, the distance $a = \cos\varphi \approx 0.60$ and $b = \sin\varphi \approx 0.80$. It follows that $a/d = 1/5$.

Because of their internal point-like charges, water molecules are tripoles, but it is customary to replace them by dipoles. They are constituted by the central charge $-2q$ and a single charge $+2q$, situated in the middle between the charges $+q$. This dipole is represented in **Figure 2** by a red arrow. By definition, the dipole moment is then $p = 2qa$. The limited validity of this approximation appears when we calculate the electrostatic potential $V(r, \theta)$ at large distances r from the center of the effective dipole. The angle θ specifies the chosen direction with respect to the axis of the dipole.

The test charge +1 does then "see" the charges along parallel lines, but their distances are slightly different. They are indicated by thin red lines. Adopting also natural units ($4\pi e^2/\delta\varepsilon_o = 1$) for electrostatic potentials, their sum is

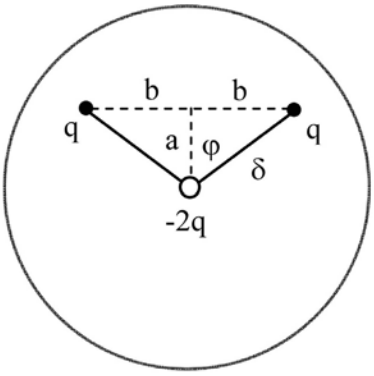


Figure 1. Model of water molecules.

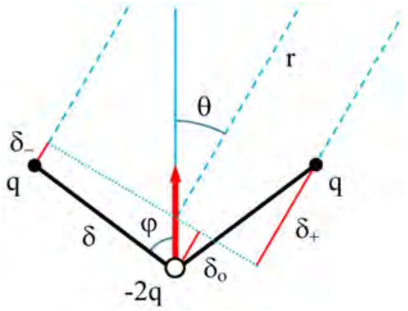


Figure 2. The dipole approximation.

$$V(r,\theta)=\frac{-2q}{r+\delta_o}+\frac{q}{r-\delta_+}+\frac{q}{r+\delta_-}$$

$$=\frac{q}{r^2}\left(2\delta_o+\delta_+-\delta_-\right)+\frac{q}{r^3}\left(-2\delta_o^2+\delta_+^2+\delta_-^2\right)+\cdots$$

The first and second order approximations result from $(1+x)^{-1}=1-x+x^2$ when $x\ll 1$. Since $\delta_o+\delta_+=\delta\cos(\varphi-\theta)$ and $\delta_o-\delta_-=\delta\cos(\varphi+\theta)$, the value of $2\delta_o+\delta_--\delta_+=2\delta\cos\varphi\cos\theta$, where $\delta\cos\varphi=a$. Thus,

$$V(r,\theta)\approx\frac{p}{r^2}\cos\theta,\quad E_r\approx\frac{2p}{r^3}\cos\theta,\quad E_\theta\approx\frac{-p}{r^3}\sin\theta \tag{1}$$

The radial and angular components of the electric field ***E*** at the observation point result from partial derivatives of $V(r,\theta)$. At closer range, there are inevitable corrections. In liquid water, it is also necessary to account for intermediate water molecules, since they are easily reoriented by an applied electric field. The potential $V(r,\theta)$ is then reduced by the factor $1/\varepsilon_r$, where the relative dielectric constant $\varepsilon_r\approx 80$. When neighboring water molecules are subjected to an electric field, their effective dipoles will be aligned. These *molecular chains* are broken by thermal agitation when the electric field is extinguished, but they could also be stabilized by association, like sticks in a bundle.

This possibility merits attention, since it is known that electric field lines can be visualized by means of neutral particles, like grains of semolina or short plastic filaments dispersed on oil. The applied electric field does merely polarize these particles, but the induced dipoles tend then to align one another. Could similar chains be formed by means of water molecules? To answer this question, we have to examine all possible types of interactions between water molecules.

2.2. The Origin of Van der Waals Forces

Even electrically neutral molecules attract one another in gases, because of Coulomb forces and quantum-mechanical effects. Indeed, if such a particle were subjected to an electric field, it would displace all weakly bound electrons inside this particle. This produces surface charges that create a secondary electric field inside the particle. It opposes displacements of the electrons and would restore neutrality when the applied field is switched off. For oscillating electric fields, this force leads to a resonance for the ensemble of oscillating electrons. That explains the appearance of colors and peculiar optical properties of thin granular metal films. They were said to be "anomalous" until they could be explained in terms of collective oscillations of nearly free electrons [21]. Since oscillations of electrons inside neutral particles do also create an oscillating electric field outside these particles, two neutral ones can interact with one another.

Nearly free electrons will be set in coupled oscillations inside neighboring particles. Their resonance frequency is then reduced, but in quantum mechanics (QM), the lowest possible energy of an oscillator is proportional to its resonance frequency. The (zero-point) energy of two neutral particles is thus reduced when they come close enough to one another. This effect can be interpreted as resulting from an attractive force. The existence of this short-range force was discovered by *Van der Waals* in 1873, since a dense gas does not behave like an ideal one. It corresponds to a model, where velocities are only randomized by collisions of point-like particles, but neutral particles attract already one another at some small distance. The physical origin of this force could only be understood after the development of QM.

Since Van der Waals forces are proportional to the volume of the polarizable particles, they are negligible with respect to other forces for water molecules in the liquid state. However, small metal particles that are suspended in liquid water contain nearly free electrons. They are very polarizable and big enough to attract one another by Van der Waals forces. In liquid water, small metal particles attract thus one another and constitute chains. These "necklaces of pearls" are observable by optical microscopy [22] and attract now much attention, because of expected applications. Similar chains might be relevant for water memory.

2.3. Hydrogen Bonds and Exchange Effects

The concept of so-called "hydrogen bonds" was already introduced in 1920, since some quantum effects could be treated in a semi-classical way [23], but simplified models can lead to confusions. The Lewis model, proposed in 1916, was merely based on the fact that many atoms are more stable when their shells contain 2 or 8 electrons. Because of Bohr's semi-classical model of atomic structures and Pauli's exclusion principle, these values correspond to closed shells. Hydrogen atoms contain only 1 electron, while C, N, O, and F atoms do respectively have 4, 5, 6 and 7 electrons in their external shells instead of 8. Molecules like CH_4, NH_3, OH_2 and FH would thus result from the "tendency to complete… the octet of electrons". This lowers the total energy, but H_2O molecules are special. The left part of **Figure 3** represents closed shells of the oxygen atom and the hydrogen atoms by blue rings. The shared electrons are represented by dots, but the oxygen atom is then surrounded by 4 pairs of electrons. There are 2 bound pairs and 2 free pairs.

It was therefore proposed that the negative charge density of the free pairs "might be able to exert sufficient force" on two neighboring oxygen atoms. This would account for mutual attraction of water molecules that allows for structuring of liquid and frozen water. The right part of **Figure 3** represents these

Figure 3. Semi-classical concepts of hydrogen bonds.

"hydrogen bonds" by means of red lines, when neighboring water molecules are assumed to be situated in the same plane. QM revealed that electrons behave according to laws that apply to waves. Every oxygen atom contains two strongly bound electrons and four external electrons in ($2s^1 2p^3$) states. Superposition of these wave functions leads to interference effects and the charge distribution of the 4 external electrons acquires then tetragonal symmetry. We might thus think that hydrogen bonds are merely due to stronger electrostatic attraction, but modern biochemistry states that "in a hydrogen bond, a hydrogen atom is shared by two other atoms" [24]. The right part of **Figure 3** can then be interpreted in terms of two donor sites of protons and two acceptor sites. This description implies that the intermediate proton might change its position. Hydrogen bonds would then be due to *exchange effects*, which are also known for nuclear forces.

Figure 4(a) represents the effects of hydrogen bonds between water molecules in 3D space. The red dots define average positions of the cores of neighboring water molecules in liquid water, while the red lines correspond in a schematic way to electron pairs, but also to possible exchanges of protons between pairs of oxygen atoms. This configuration requires a modification of the internal structure of water molecules, since the normal angular separation of two protons was there $2\varphi = 106^\circ$ (**Figure 1**). Here we get 4 equal angles $2\mu \approx 110^\circ$. They are defined by joining the center of the cubic cell to four equally separated vertices. The value of $\cos\mu = h/c = 1/\sqrt{3}$, since $(2c)^2 = 3(2h)^2$. It follows that $\mu = 54.74^\circ$.

Figure 4(a) accounts only for one possible lattice structure of ice, since imposed temperatures and pressures can yield different phases for frozen water [25]. In liquid water, adjacent molecules are moving around, since all bonds are constantly broken and reconstituted at an extremely rapid pace [26]. Nevertheless, small-scale order is statistically preserved, while large-scale order is lost. At an intermediate scale, we get the extended lattice-structure of **Figure 4(b)**. Body-centered and empty cubes are alternating. This yields many voids, which will often be filled at higher temperatures. This fact explains the existence of low-density and high density liquid water, as well as analogous amorphous states for frozen water.

It is important to be aware of the quantum-mechanical nature of exchange effects. They are also possible between two X and Y atoms, when an intermediate proton could belong to X or Y. Both possibilities are expressed by the notation XH… Y or X… HY. QM accounts indeed for limited knowledge. The probability distribution for possible positions of *electrons* is defined by means of their wave functions. Exchange effects are then due to "tunneling" through an intermediate

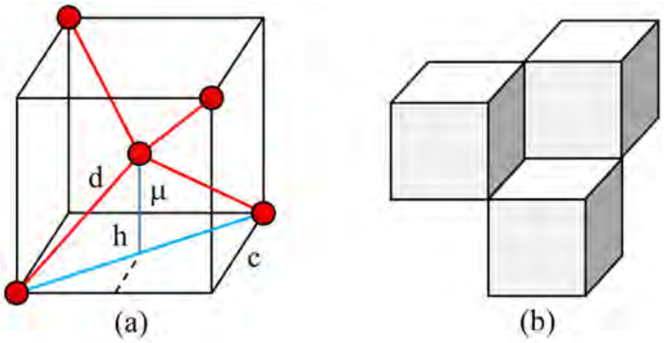


Figure 4. (a) Idealized model for the relative positions of oxygen atoms in neighboring molecules for ice and liquid water; (b) Extended lattice structure of this type.

potential barrier. This is relevant for bonds between atoms inside molecules and in particular for H_2^+, where one electron allows for H-H$^+$ or H$^+$-H. Although protons have a greater mass than electrons, they are also subjected to quantum-mechanical laws. The probability distribution for being at different places in space is then not smeared out, but reduced to needle-like (delta) functions. When a proton has two possible positions, it can be said to be *delocalized,* but possible exchanges are then not due to tunneling. They result from the fact that the energy of any physical system cannot be precisely determined during short time intervals Δt. There is always an irreducible uncertainty $\Delta E \approx h/\Delta t$. In semi-classical terms, the proton is able to "jump" over the intermediate potential barrier, when this happens rapidly enough.

2.4. Coulomb Forces and Exchange Effects for Water Dimers

To analyze possible effects of protons for interacting water molecules, we begin with the simplest case. Figure 5 represents the cores of two oxygen atoms by open dots, separated by the distance d. The left part of this figure corresponds to purely classical concepts. The intermediate proton has a well-defined position, indicated by a black dot. Since the measured length of OH bonds is δ, we can calculate the total potential energy U_1 that results from electrostatic interactions between two point-like particles of charge −2 and one charge +1. However, the concept of *hydrogen bonds* means that the intermediate proton has two possible positions. They are represented by gray dots in the upper right part of Figure 5. When they are occupied with equal probabilities, we have to attribute an average charge +1/2 to these positions. However, even when a particle is delocalized, it cannot exert forces on itself. We will calculate the resulting electrostatic potential energy U_2. The third configuration would be obtained if it were possible to account for hydrogen bonds in a semi-classical way, by assuming that the proton has only one well-defined position, situated in the middle. This position is represented by a black dot and the potential energy would then be U_3.

Using natural units for charges, distances and energies, we get

$$U_1 = \frac{4}{d} - \frac{2}{\delta} - \frac{2}{d-\delta} = -1.67 = U_2 \quad \text{and} \quad U_3 = \frac{4}{d} - \frac{2}{d} = -0.69$$

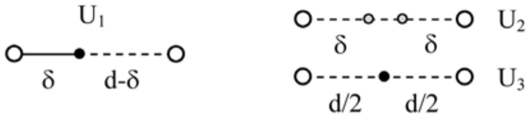


Figure 5. Three conceivable models for O-H-O bonds.

Since $U_2 = U_1$, a strictly classical description is sufficient, although the proton is delocalized. This would even be true (in the present case) if the uncertainty did allow for any partition q' and $(1 - q')$ of the charge +1. The third configuration is very unstable and does not account for quantum-mechanical exchange effects. We are now ready to calculate the total potential energies V_1 to V_4 for different configurations of dimers, represented in Figure 6. They result from the fact that two water molecules can easily be rotated with respect to one another, but this modifies the potential energy of the interacting tripoles. The usually assumed configuration of $(H_2O)_2$ is close to the upper left one of Figure 6. The tripole of one molecule is situated in the plane of the drawing and one of its two protons is precisely oriented towards the core of the neighboring molecule. The other tripole is perpendicular to this plane and seen in profile. We will calculate the potential energy V_1 for the indicated configuration.

V_2 corresponds to aligned effective dipoles. This configuration would be preferred if water molecules did only contain dipoles, but could even be privileged for effective dipoles, when the dimer is subjected to an electric field. We want thus to see if V_2 is already close to V_1 in the absence of an applied electric field. The tripoles should then be orthogonal to one another to minimize repulsion between protons in neighboring molecules. V_3 is the potential energy for any pair of molecules in a nearly linear chain, where all intermediate protons are ideally situated between two oxygen atoms. The resulting zigzagging configuration is planar and in conformity with a classical representation.

An applied electric field could even allow for a perfectly linear chain, because of *intramolecular* exchange effects, although this was unknown. For clarity, we consider here two coplanar tripoles. One proton is always situated as close as possible to the core of the neighboring oxygen atom. The other proton has two equally probable positions, above and below the symmetry axis. This would yield the energy V_4 for any pair of water molecules. Alternatively orthogonal tripoles would reduce repulsion between the delocalized protons. The resulting potential energy is then $V_5 < V_4$.

To facilitate this type of calculations, we note that the total potential energies result always from adding the Coulomb potentials ($V = qq'/\Delta$) for pairs of point-like charges q and q', separated by a distance Δ. We define thus a function $S(x, y, z)$, where x, y and z are differences of Cartesian coordinates, respectively measured towards the right, rear and top. Thus,

$$V = qq'S(x,y,z) \quad \text{where} \quad S(x,y,z) = \left(x^2 + y^2 + z^2\right)^{-1/2} \tag{2}$$

The values of V_1 depends on the angle $\varphi = 106^\circ$ and the complementary angle $\phi = 74^\circ$. Since $\delta = 1$ in natural units, the distance $a_1 = \cos\phi = 0.28$ and $b_1 = \sin\phi$

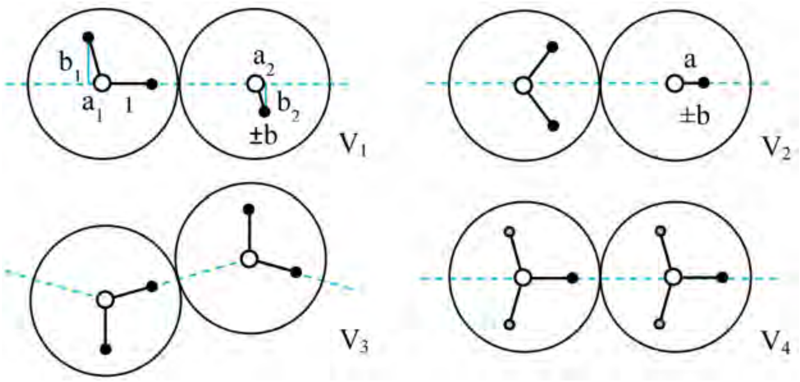


Figure 6. Conceivable structures of water dimers.

= 0.96. However, $a_2 = aa_1$ and $b_2 = ab_1$, where $a = \cos\varphi = 0.60$ and $b = \sin\varphi = 0.80$. Thus,

$$\begin{aligned} V_1 &= 4S(d,0,0)+2S(d-1+a_2,b,b_2)+2S(d+a_1+a_2,b,b_1+b_2) \\ &\quad -2S(d-1,0,0)-4S(d+a_2,0,b_1)-2S(d+a_1,0,b_1) \\ &= -0.116 \end{aligned}$$

$$V_2 = 4S(d,0,0)+4S(d,b,b)-4D(d+a,b,0)-4D(d-a,0,b) = -0.093$$

V_1 can be slightly lowered when the repulsion between the nearest protons is reduced by a small rotation of the left molecule around its center. An angle of 1.4° is sufficient to reach the minimal potential energy [27]. Since $V_1 < V_2$, it has been assumed that more than 2 water molecules should always be assembled according to the same rule as for the most strongly bound dimer. Water molecules could then only constitute rings or clusters of limited size, but *linear polymerization* is not excluded for water molecules. It would even be preferred for n_2 water molecules, compared to clusters of n_1 molecules, when $(n_2 - 1)V_2 < (n_1 - 1)V_1$. It is thus sufficient that $n_2 > 1.3n_1$. Of course, long chains of water molecules would be too fragile to resist thermal agitation in liquid water, but we will show (in Section 2.7) that chains of water molecules with aligned effective dipoles can be stabilized.

Moreover, $V_3 = -0.111$, which is quite close to $V_1 = -0.116$. Since water molecules can easily be rotated, they are aligned according to this pattern inside very narrow pores [28]. They are then said to form wires or filaments. Perfectly linear polymerization with intramolecular exchange effects would yield

$$\begin{aligned} V_5 &= 5S(d,0,0)+S(d,b_1,b_1)+S(d-1-a_1,b_1,0)+S(d+1+a_1,0,b_1) \\ &\quad -2S(d-1,0,0)-2S(d+1,0,0)-2S(d-a_1,b_1,0)-2S(d+a_1,0,b_1) \\ &= -0.084 \end{aligned}$$

An applied electric field does not only align water molecules, but also polarize water molecules in such a chain. This yields stronger bonds, since all positive charges of one molecule come closer to the central negative charge of the neighboring molecule. Polarization of water molecules is possible [29] and favors thus linear polarization by means of intramolecular exchange effects.

2.5. Initial Evidence of Molecular Chains

When the young *Theodor von Grotthuss* was experimenting in1806 with a Volta pile, he discovered that pure water has a much higher electric conductivity than other liquids. The chemical structure of water molecules was not yet known. [John Dalton asserted in 1808 that it is HO, while Avogadro proposed in 1811 that it could be H_2O. This hypothesis was disregarded, since it resulted from the assumption that all types of particles occupy the same volume in gases, whether they are molecules or atoms. This was only accepted at about 1860, after the development of the kinetic theory of gases.] However, Von Grotthuss knew that water molecules are composed of positive and negative parts, since they can be separated by electrolysis. This had already been proven in 1800.

Von Grotthuss thought therefore that water molecules are held together in liquid water by mutual attraction of positive and negative parts. An electric field should align them. The high electric conductivity of liquid water could then be explained, if water molecules did contain tiny charge carriers that move more easily inside these chains than outside them [30]. This hypothesis may have been suggested by the method of fire-fighting of that time. People were standing in a row and passed buckets from hand to hand, but the existence of protons was not yet known. Nevertheless, this explanation was appealing and successful. Now, we can justify this model in term of *intramolecular* exchange effects. The left part of **Figure 7** represents two water molecules in a linear chain, subjected to an electric field $\boldsymbol{E}$. One molecule contains an additional proton. We indicate displacements of charged particles by red lines, but the applied electric field does also modify the intermediate potential barrier. It becomes dissymmetric and facilitates "hoping" over the potential barrier for the proton in best position. The blue arrow represents a jump of this proton towards one of the nearby potential wells.

The right part of **Figure 7** shows the result (without polarization effects). The H_3O^+ ion became a normal H_2O molecule, while the neighboring H_2O molecule was converted into an H_3O^+ ion. The proton did advance without being deviated by collisions. [Ohm's law is still valid for relatively small electric fields, as for ionic conduction in solid state physics]. We could equally well consider H_2O molecules with one delocalized proton on the left side and an adjacent HO^- ion. It contains only one proton, attracted toward the center of the next molecule. The electric field $\boldsymbol{E}$ would then cause a jump of the intermediate proton towards one of the two empty places. This would be equivalent to opposite motion of a proton hole. The essential result is that *intramolecular* exchange effects are realistic, since they explain the high electric conductivity of liquid water in a more detailed way.

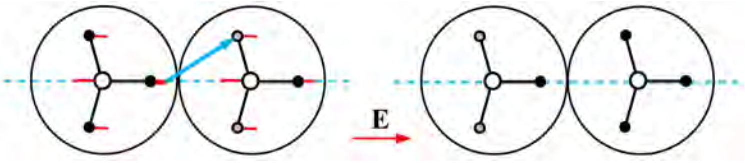


Figure 7. Explanation of the von Grotthuss mechanism.

2.6. Evidence of 2D Polymerization of Water Molecules

The story of this discovery is similar to that of water memory. It began with an unexpected observation, made by the Russian chemist *Nikolai Fedyakin.* He condensed water in thin capillary quartz tubes and found that its physical properties are different from those of ordinary water. This phenomenon was totally unexpected, but arose at first much attention and scientific curiosity. *Lippincot and Stromberg* combined, for instance, routine infra-red spectroscopy with an improved method for producing this type of water. They confirmed that it has peculiar properties and proposed an explanation [31]. They assumed that water molecules can be bound to one another inside layers like that of **Figure 8**. The nature of hydrogen bonds was misunderstood (see Section 2.4), but this configuration was simple and heuristically useful. Since the layers can slide on one another, water remains liquid, but is viscous near substrates that favor this configuration.

2D polymerization of water molecules would thus yield a "new state" of liquid water, since it is partially crystallized. However, other persons declared that this is impossible, since they preferred to stick to customary ideas. The media propagated the slogan of "bad science", which had great impact. It led even to total prohibition of research on "polywater". Academic careers would have been broken for anyone who might dare to be involved in such "pathological" science. Even Stromberg, interviewed some 40 years later, accepted that researchers might be misled by unconscious bias [32]. However, he added that "most mistaken hypotheses in science aren't entirely wrong; they just have to be modified a bit."

Actually, it is well-known in material science that crystallization can be influenced by the substrate, because of local attractions. *Rostrum Ray* proposed that this phenomenon of epitaxy might explain water memory, because of the "extreme structural flexibility" of water molecules [33]. Their internal structure and relative positions could be modified, but this model would require molding and stability of detached structures and even multiplication of positive and negative molds in the course of successive EHDs. These hypotheses are not plausible enough, but Ray tried at least to explain water memory, instead of simply pretending that it is impossible.

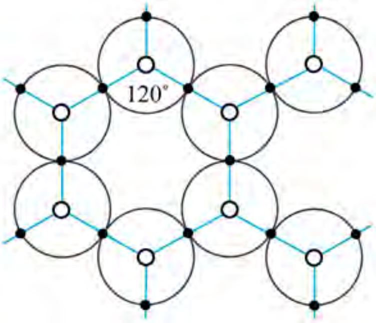


Figure 8. Polywater with hexagonal cells.

2D polymerization of water molecules was rediscovered by the bioengineer *Gerald Pollack*. He wondered why particles that imitate red blood cells can easily move through narrow capillaries. Trying to understand this fact, he realized that some materials create an "exclusion zone" near their surface [34]. Indeed, water molecules can be so strongly bound to the substrate and to one another in successive layers that the presence of red blood cells, for instance, becomes there impossible. Pollack adopted the model of **Figure 8** to explain the formation of 2D lattices. He deduced from this model that every oxygen atom is surrounded by 3 half-hydrogen atoms. In natural units, this would yield a charge $(3/2) - 2 = -1/2$ per molecule. Exclusion zones should thus be electrically charged and this was proven to be true. Contact of pure water with some materials is sufficient to constitute a battery.

Pollack's discovery and empirical investigations were outstanding achievements, but **Figure 8** can be replaced by **Figure 9**. Delocalized protons have then 3 equally probable positions inside any water molecule. This model combines intramolecular with intermolecular exchange effects. It is then not necessary to assume strong distortions of tripoles in all water molecules, since the normal angles of 106˚ can be preserved when the third angle is 148˚. Regular hexagons would merely be transformed into elongated ones. The upper and lower rows of **Figure 9** are even zigzagging chains, like those of **Figure 6**.

2.7. Formation and Stabilization of Molecular Chains

In interstellar space, there are ions that attract water molecules and align their effective dipoles, since the configuration V_2 of **Figure 6** is sufficient for moderate electric fields. **Figure 10** represents such a molecular chain in 3D. The tripoles are alternatively orthogonal to one another and the aligned effective dipoles are represented by red arrows. On the average, the chain is axially symmetric. In outer space, ions would thus collect water molecules and become radially "haired". This does facilitate the participation of water molecules in the formation of planetary systems.

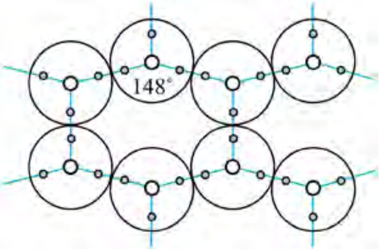


Figure 9. Polywater with elongated cells.

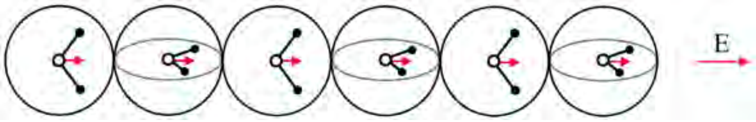


Figure 10. A single chain of water molecules.

Although positive and negative ions are strongly bound to one another in ionic crystals, they are easily dissolved in liquid water, since the small water molecules are more attracted. They penetrate inside ionic crystals and dissolve them. In liquid water, ions will thus usually be isolated and surrounded by a "hydration sphere", where the effective dipoles of water molecules are oriented towards the central charge. This polarization decreases further away, because of thermal agitation. However, molecular chains could also be formed and rapidly stabilized by attracting one another. **Figure 11** represents two possibilities by means of three water molecules that belong to parallel chains.

According to the dipole approximation (1), the electric field is $-p/r^3$, when $\theta = 90^\circ$. Real dipoles would thus be antiparallel in lateral positions, but we have to consider chains of tripoles. Two parallel molecular chains will thus be shifted by the distance a with respect to one another. This allows for parallel *or* antiparallel effective dipoles. We expect that parallel ones are preferred, since the two protons of the lower molecule are then closer to the negative center of the upper right molecule. However, it is useful to verify if this leads to significant differences for the resulting potential energies V_a and V_b. Since the potential energy of the upper pair is $V_2 = -0.093$ in natural units, we get

$$\begin{aligned} V_a &= V_2 + 4S(a,0,d_a) + 2S(a,b,d_a \pm b) + 4S(D-a,0,d_a) \\ &\quad + 2S(d_+a,0,d_a) + 2S(d_+a,2b,d_a) - 2S(0,0,d_a \pm b) \\ &\quad - 4S(2a,b,d_a) - 4S(d,b,d_a) - 4S(d-2a,b,d_a) \\ &= -0.28 \end{aligned}$$

$$\begin{aligned} V_b &= V_2 + 4S(a,0,d_a) + 2S(a,b,d_a \pm b) + 4S(d-a,0,d_a) \\ &\quad + 2S(d+a,2b,d_a) + 2S(d+a,0,d_a) - 2S(0,0,d_a \pm b) \\ &\quad - 4S(2a,\ b,\ d_a) - 8S(d,b,d_a) \\ &= -0.09 \end{aligned}$$

The ± signs mean here that we have to sum two different terms. It appears that V_a is 3 times lower than V_b. Agglomerations of *parallel* chains of water molecules are thus preferred. Moreover, the effective dipoles are already oriented in nearly the same way by the electric field of the ion. Molecular chains with parallel effective dipoles get spontaneously assembled and stabilized.

It is very important to realize that biologically active molecules contain electrically charged parts. They explain why these molecules are easily dissolved in liquid water and require at least contact with saliva. **Figure 12** shows some typical examples. We see that the charged parts are even situated on a branch, where

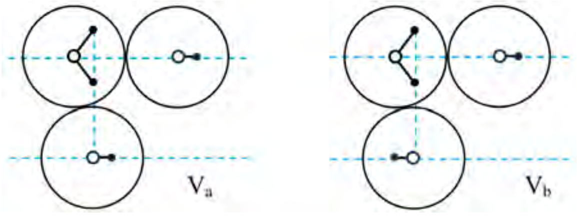


Figure 11. Mutual attraction of molecules on parallel chains.

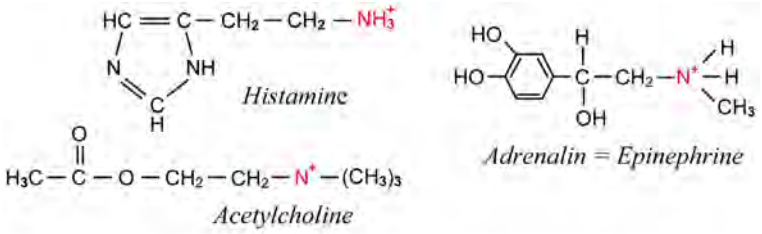


Figure 12. Biologically active molecules contain charged parts that can vibrate.

they can be set in oscillation by thermal agitation of water molecules in the surrounding liquid. They *resonate* at a particular frequency, which is characteristic of the active molecule. Its value is much lower than for vibrations of strongly bound charged particles inside molecules of any type. For active molecules, the resonance frequency is determined by the effective mass of its charged part and a weak restoring force. It results from deformations of the soft cocoon of polarized water molecules. These ideas are essential to unravel the puzzle of water memory.

2.8. Single-Domain Ferroelectric Crystallites

When the electric field of a biologically active molecule has started to assemble water molecules, it becomes a germ of ongoing crystallization. More and more water molecules are attracted and align their effective dipoles. This yields closely packed molecular chains, like that of **Figure 10**. A frontal view of the alternatively orthogonal tripoles yields a square lattice. Ideal alignment of all effective dipoles is only achieved in "single-domain ferroelectric crystallites". Water molecules are there more closely packed than in the surrounding water. We will prove that these crystallites contain many water molecules and are thus spherical. We will call them *water pearls* (WPs) and prove that they have the same size. Since they contain a great number N of water molecules, their equally oriented effective dipole moment p yields a very great total dipole moment $P = Np$ for every WP. Each one of them creates thus an electric field, which is able to assemble other water molecules. This process creates an adjacent WP and even a spontaneously growing chain of WPs.

To discard unnecessary objections, we mention that the solid state physicist *Kittel* realized already in 1946 that molecules of magnetite (Fe_3O_4) create single-domain ferromagnetic crystallites [35]. Since they are ideal magnets, their discovery led to important applications, like magnetic recording. Blackmore discovered in 1975 that some species of aquatic bacteria collect Fe_3O_4 or Fe_3S_4 molecules. They are then spontaneously assembled and constitute single-domain ferromagnetic crystallites [36]. They form chains of about a dozen beads. Magnetotactic bacteria produce these chains to remain in deeper layers of shallow waters and are able to sense the orientation of their internal compass needles in the inclined geomagnetic field of the Earth. They developed this stratagem to survive, since the oxygen content would be too high for them near the surface. This required the formation of specialized genes [37].

To determine the radius R of these ferroelectric crystallites is a tricky problem. It has been tackled for the most common ferroelectric material ($BaTiO_3$) by means of the theory of phase transitions [38]. Since this approach does not apply to water molecules in the liquid state, we adopt another one. It does simply result from a transposition of the method that is used in electrostatics to determine the polarization of a homogeneous spherical particle. Figure 13(a) represents such a WP in pale blue color. Because of its surface charge, it orients water molecules in the surrounding liquid, but only in a limited region. The polarization of the surrounding water molecules is there progressively decreasing because of thermal agitation. This region is represented in a different color.

We use polar coordinates (r, θ) with axial symmetry. On the average, the WP is electrically neutral, because of the closeness of the charges $\pm 2q$ inside all water molecules. The internal surface charge density results from the charges $\pm 2q$ at the extremities of every molecular chain. It occupies the surface d^2 in planes that are perpendicular to these chains, but the surface of the sphere is inclined. The intersected surface is thus increased and the internal surface charge is $\sigma_i(\theta) = (2q/d^2)\cos\theta$ at the positive side. The external surface charge density $\sigma_e(\theta)$ is lower, but proportional to the internal one. The total surface charge density is thus $\sigma(\theta) = \sigma_o\cos\theta$. Positive and negative surface charges on opposite sides of a WP create a homogeneous electric field E_i inside this sphere, as if it were composed of many very thin condensers. The electrostatic potential inside the WP is therefore $\phi_i(r, \theta) = E_i r\cos\theta$.

Figure 13(b) defines the potential $\phi(r) = (Q/r)\exp(-r/\lambda_o)$ that would be produced outside the sphere by the total charge $Q = N2q$ of all effective positive poles of water molecules, if this charge were situated at the center of the sphere. The usual Coulomb potential (Q/r) is modified by *screening effects,* resulting from positive and negative charges in the polarized region of the liquid water. The radial decrease is characterized by the *Debye length* λ_o. Figure 13(c) defines the actual potential $\phi_e(r, \theta)$ in the external medium. It is due to the charges $\pm Q$ of all positive poles and negative poles, separated by the distance a. The external potential depends then on $\Delta r = (a/2)\cos\theta$, since

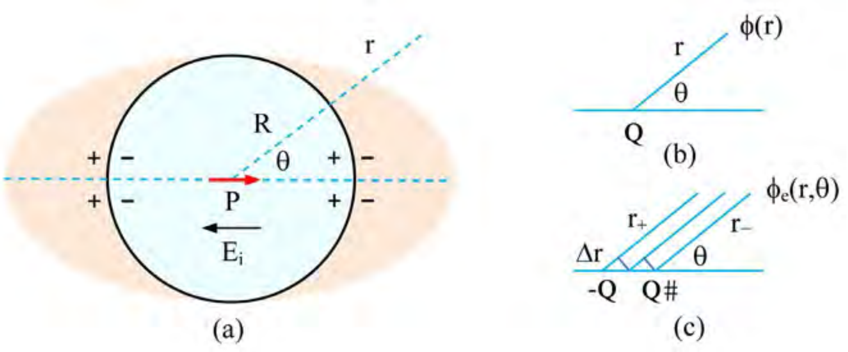


Figure 13. (a) A water pearl and the external domain of oriented water molecules; (b) A single charge Q situated at the center of the sphere would generate an electric potential $\phi(r)$. (c) The actual external potential $\phi_e(r, \theta)$ results from two charges $\pm Q$.

$$\phi_e(r,\theta) = \phi(r-\Delta r) - \phi(r+\Delta r) \quad \text{for} \quad r \geq R$$

ϕ_i(r, θ) and ϕ_e(r, θ) are subjected to boundary conditions, which determine the values of R and E_i. This more technical problem is solved in the short appendix, but all required physical concepts have been explained here and the result is that $R \approx 10\lambda_o$. The value of the Debye length λ_o depends on the concentration of ions in the surrounding water. It is also proportional to the square root of the absolute temperature T, but this factor is nearly constant between 20 and 30˚C. The value of λ_o has been measured at 25˚C for water with different concentrations of dissolved NaCl [39]. It appeared that $\lambda_o \approx 1$ nm at 0.1 M (mol/liter), but increases for lower and higher concentrations. Actually, $\lambda_o \approx 3$ nm at 0.01 M and 5 M, which is only slightly higher than for Dead Sea water. For pure water, λ_o would depend on the concentration of H^+ and OH^- ions. The radius R of WPs is then somewhat smaller than 10 nm.

2.9. Properties of WPs and Their Poles

When the particle physicist *Shui-Yin Lo* was visiting professor at the famous California Institute of Technology in 1996, he adopted a research project concerning properties of liquid water. He was surprised to discover that EHDs of HCl, NaOH or HNO_3 molecules in very pure water led to the formation of "novel stable structures". Lo thought that they result from crystallization of hydration spheres [40], while we attribute their existence to the formation of ferroelectric crystallites. This explains the dipolar nature of these particles and their great stability. They subsisted when all ions had been removed by successive dilutions and these structures were even multiplied in the course of successive dilutions, but only when they were followed by vigorous agitation.

When S. Y. Lo determined the sizes of various types of structures by means self-interference of scattered laser light, he found 3 distinct groups. The smallest particles had a diameter $D \approx 15$ nm with very low dispersion. We interpret this result as meaning that D is the diameter of WPs in pure water. Thus, $R \approx 7.5$ nm and the Debye length $\lambda_o \approx 0.75$ nm. Since the volume occupied by every water molecule is d^3, where d = 0.275 nm, WPs contain $N \approx 85{,}000$ *molecules*. This huge number justifies the assumption that they are spherical. Nevertheless, WPs are nanoparticles, since water molecules are very small.

The second group of structures, discovered by S.Y. Lo, had a size of about 300 nm. We consider that this group corresponds to the length $L = ND$ of chains of WPs, containing $N \approx 20$ water pearls. We will explain (in Section 3.3), why their length L has to be limited. Its value depends on the mutual attraction between positive and negative surface charges on adjacent hemispheres. It is thus useful to replace the distributed surface charges of WPs by point-like poles. They are situated inside the sphere, like those of magnetized steel balls, but we can be more explicit. The total charge Q_o on the surface of the positive hemisphere, is the integral of $2\pi r\sigma(\theta) r\mathrm{d}\theta$, where $r = R\sin(\theta)$ and $\sigma(\theta) = \sigma_o\cos\theta$, while the angle θ varies from −π/2 to + π/2. This yields $Q_o = (4\pi/3)R^2\sigma_o$. We can also calculate the electrostatic potential V(x) for a test charge +1 that is situated on the symmetry axis at the distance *x* from the center O of a WP. **Figure 14** shows that V(x) is the integral of $2\pi\sigma(\theta)r^2\mathrm{d}\theta$ divided by the distance *R'*. This allows us to define the effective charge Q(x) of the positive pole, if it were situated at the distance *R*/2 from the center O. This pole is represented by a black dot and

$$V(x)=\frac{3Q_o}{2}\int_{-\pi/2}^{+\pi/2}\frac{\sin^2\theta\cos\theta\mathrm{d}\theta}{\left[R^2+x^2-2Rx\cos(\theta)\right]^{0.5}}=\frac{Q(x)}{x-(R/2)}$$

The result of numerical integration is shown in **Figure 15**. It appears that Q(x) = Q_o when $x > 3R$, but when the test charge is close to the surface of the WP, it does mainly interact with the nearest surface charges. This reduces the value of Q(x). On the surface, Q(R) = $0.68Q_o$. We neglected all screening effects, but it is only important that neighboring poles are separated by the same distance *R* and carry charges $\pm Q$.

3. The Mechanism of Water Memory

3.1. Small Oscillatory Rotations of Water Pearls

Because of the rapidly decreasing Coulomb forces, it is sufficient to consider the mutual attraction of neighboring positive and negative poles. At rest, they are aligned and their poles are separated everywhere by the same distance *R*, but small oscillatory rotations of WPs around their center will lead to transverse displacements of the poles. They are represented in **Figure 16** by red arrows. For clarity, we exaggerated their magnitude. The essential point is, indeed, that neighboring poles remain in tangential contact and that the transverse displacements of the poles can vary along the chain. We characterize the instantaneous rotation of the n^th^ water pearl by the variable $u_n(t)$.

Figure 17 shows that when u is the relative displacement of two neighboring poles, they attract one another by the force ***F***. In natural units, its magnitude is

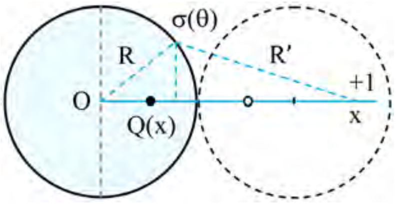


Figure 14. Definition of the charge *Q*(*x*).

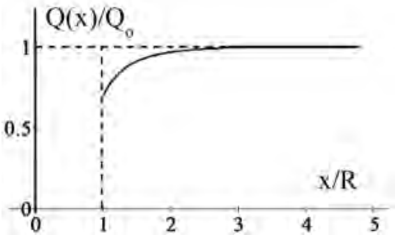


Figure 15. Calculated value of *Q*(*x*).

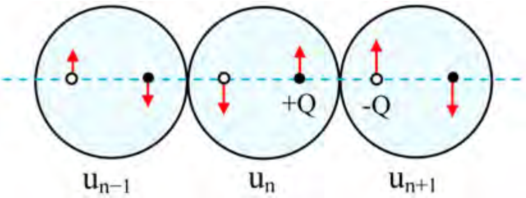


Figure 16. Small rotations of water pearls in a chain are specified by displacements of poles.

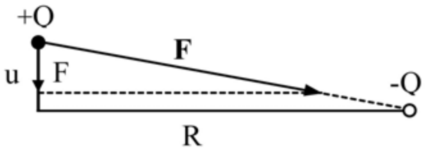


Figure 17. The restoring force is the transverse component of the force *F*.

$Q^2/(u^2 + R^2)$. The transverse component F is reduced by the factor u/R. It follows that when $u \ll R$, the *restoring force* is

$$F(u) = -Ku \quad \text{where} \quad K = \frac{Q^2}{R^3} \tag{5}$$

This force is proportional to the relative displacement u, as for any elastic system. When M is the effective inertial mass of poles, the equation of motion for the n$^{\text{th}}$ water pearl is

$$M\ddot{u}_n = K(u_{n+1} - u_n) - K(u_n - u_{n-1}) \tag{6}$$

Every dot stands for a time derivative. Since this equation is identical for all WPs, it describes the behavior of the whole chain. It can be simplified when the displacements u_n are smoothly varying along the chain, which is equivalent to saying that the diameter D of WPs is small compared to the distance where the relative displacements u_n are notably varying. We can then replace $u_n(t)$ by $u(x, t)$, where the coordinate x is treated as if it were a continuous variable. Actually, $u_{n\pm1} = u_n \pm Du' + D^2u''/2$, where u' and u'' designate first and second order partial derivatives with respect to x. Equation (6) is then reduced to

$$\ddot{u} = v^2 u'' \quad \text{where} \quad v^2 = D^2K/M \tag{7}$$

This is the usual wave equation for vibrating strings. An infinite chain would allow for $u(x,t) = u(x \pm vt)$. This corresponds to a function of any shape, moving at the velocity v towards the right or the left. Possible attenuations of oscillatory rotations have been neglected in (6) and (7), but will be discussed later on.

3.2. Standing Waves on Chains of Water Pearls

For a chain of finite length L, we have to know the boundary conditions at $x = 0$ and $x = L$. When both ends are free, the first and last pearls are not subjected to any force. Thus, $u'(x,t) = 0$ for $x = 0$ and $x = L$. In other words, $u(x, t)$ has to reach maximal values at both extremities. This allows for a particular solution of well-defined frequency f and well-defined wavelength λ:

$$u(x,t) = A\cos(kx)\sin(\omega t) \quad \text{where} \quad \omega = 2\pi f \quad \text{and} \quad k = 2\pi/\lambda \tag{8}$$

Since $u'(x,t)$ is proportional to sin(kx), the boundary condition $u'(0,t) = 0$ is satisfied, but $u'(L,t) = 0$ requires that $kL = s\pi$, where $s = 1,2,3,\cdots$. It follows that $L = s\lambda/2$ and because of (7), that the spectrum of possible frequencies is defined by

$$f = \frac{v}{\lambda} = s\frac{v}{2L} = f_s \quad \text{where} \quad s = 1,2,3,\cdots \tag{9}$$

The only possible frequencies are thus integer multiples of the fundamental frequency $f_1 = v/2L$. For sound waves, any pair of such frequencies would produce an impression of harmony. The spectrum $f_s = sf_1$ is therefore said to be a "harmonic" one. These properties are well-known in physics, but everyone should see why a chain of WPs with free ends does only allow for standing waves. This means that for any particular solution (8), all WPs are set in oscillatory rotations at the same frequency f, but the amplitude of these oscillations varies along the chain.

However, the approximation (7) is of limited validity, since it requires that $D \ll L$. The measurements of Lo imply that chains of WPs contain a relatively small number of WPs ($N \approx 20$). To see how far the approximation (9) is realistic, we have to solve the general equation (6). This is easy when we use complex notations, since standing waves are then defined by

$$u_n(t) = A\mathrm{e}^{i(kx_n - \omega t)} \quad \text{where} \quad \omega^2 = \frac{K}{M}\left(\mathrm{e}^{ikD} + \mathrm{e}^{-ikD} - 2\right)$$

Thus,

$$\omega(k) = v\frac{\sin(kD/2)}{D/2} \quad \text{and} \quad f(\lambda, D) = \frac{v}{\pi D}\sin(\pi D/\lambda) \tag{10}$$

The function $f(\lambda, D)$ is represented by the dark curve in **Figure 18**. It reveals that the domain of possible frequencies is limited, as well as the domain of possible wavelength. It is only physically possible that $\lambda/2 \geq D$. The approximation $f = v/\lambda$ corresponds to the red line, which would allow for arbitrarily high frequencies, but (10) imposes an upper limit ($f_{\max} = v/\pi D$). Nevertheless, the linear approximation is valid for a relatively large domain of low frequencies. [Indeed, $\sin(x) = x - x^3/6 \approx x$, when $x \ll 2.5$]. Actually, the spectrum of possible frequencies is $f_s = sf_1$ when $s \ll N$.

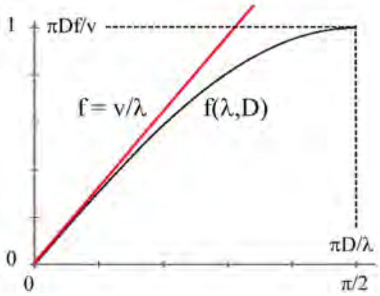


Figure 18. The general function f(λ) for chains of WPs.

3.3. Automatic Trimming of Chains of Water Pearls

Figure 19 summarizes the essential steps that explain why chains of water pearls are the *information carriers* of water memory. The upper line represents a biologically active molecule, where the charged part oscillates at some very low frequency f. It cannot communicate these oscillations to the WPs, since the length of the chain $L < \lambda/2$, where the wavelength λ is determined by $f = v/\lambda$. The second line represents what happens as soon as the length of the chain of WPs reaches the value $L = \lambda/2 = v/2f$. The electric field of the oscillating charge does then excite a standing wave for oscillatory rotations of WPs.

The last WP of the growing chain does suddenly start to oscillate when its length $L = v/2f$. The amplitude of this oscillation is the same as for the first WP of the chain, which can also be set in forced oscillation by the active molecule. It remains attached to it, but can now communicate its motion to other WPs of the chain. However, the rotation of the last pearl of the chain prevents the formation and attachment of an additional WP. The growing chain is thus *automatically trimmed.* Information that is characteristic of the type of active molecules has been encoded by means of the length L of the chain. It depends indeed on the frequency f.

The third line of **Figure 19** shows that when the trimmed chain is detached from its generator, it does still allow for a standing wave for the same length $L = \lambda/2$. The liberated chain of WPs conserves the acquired information. Moreover, it produces itself an *oscillating electric field.* Its frequency f is adequate to create more equally trimmed chains of WPs. They are thus multiplied by an autocatalytic process. This provides the key for a rational justification of water memory. It resulted from a systematic examination of all possible interactions between water molecules and their logical consequences.

We might object that oscillatory rotations of WPs will be damped by friction, exerted by surrounding water molecules. This is true, but trimmed chains are also subjected to local impacts of water molecules. Although the impacts are random, the chain does pick-up energy when it allows for resonances at any frequency f_s for possible standing waves. Since it is sufficient that the free ends of the chain can oscillate with maximal amplitude, standing waves of smaller wavelengths and higher frequencies can also be excited. Excitation of a standing wave at a higher frequency can easily be demonstrated with a flute, since "overblowing" is sufficient to double the frequency for standing waves, without modifying the effective length for longitudinal oscillations. Oscillatory rotations of WPs at

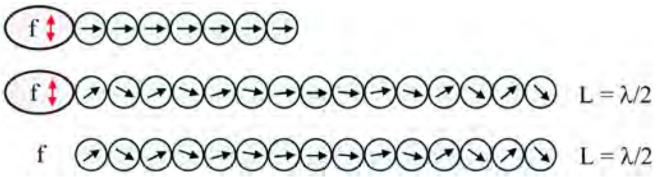


Figure 19. Creation of trimmed chains of water pearls, resonating at the frequency f of the active molecules for chains of length L. This remains true for detached chains.

higher frequencies imply more rapid motions and thus greater kinetic energies and more violent local impacts. Available energies depend on the statistical distribution of kinetic energies of water molecules in the liquid state. Although a chain of WPs of given length allows for a superposition of different modes of oscillations, those of increasingly higher frequencies will thus be excited with decreasing amplitudes. They are byproducts of random re-excitation, but the lowest possible frequency f_1 remains the predominant one.

3.4. Negation of Water Memory Was Based on False Assumptions

Benveniste's experimental proof of water memory was categorically rejected because of prevailing beliefs. They resulted from four erroneous assumptions:

1) Biologically active molecule can have no effects any more, when all of them have been eliminated by successive dilutions.

2) Even if biologically active molecules could create substitutes, made of water molecules, they would have to be adaptable. Such aggregates are unknown and can thus not exist.

3) Biologically active molecules can only act on their specific receptors by means of chemical affinities. Local structuring of liquid water would be unable to mimic these processes. This is particularly implausible for various types of molecules, since that would require an extraordinary capacity of adaptable imitation.

4) Extra high dilutions are also used for homeopathy, which is inefficient. The preparation of EHDs does even involve shaking by vigorous "successions". This ritual is a sign of charlatanism.

We have already shown that the two first objections are contradicted by the formation of trimmed chains of WPs. The third objection concerns the *fundamental* problem of molecular interactions. Modifications of the state of motion can result from direct contact (collisions), but also from actions at a distance (due to attractive or repulsive forces). We are accustomed to the idea that structural changes (combinations or dissociations) at molecular level result from chemical reactions, requiring direct contact, chemical affinities and configurational conformity. However, internal modifications can also result from energy transfer (excitation or disexcitation) by means of force fields.

Figure 20 summarizes the required restructuring of our ideas, because of the concept of ferroelectric water pearls. First of all, we have to realize that biologically active molecules contain an electrically charged part (Figure 12) that has a resonance frequency f. It creates thus an *electric field* that oscillates at this frequency. The upper line of Figure 20 represents the normal process, where this electric field acts on a molecular receptor, which has also a charged part that can oscillate. Specific receptors of a particular type of active molecules contain also a charged part that can oscillate. It resonates at a frequency f_r. It is thus sufficient that $f_r \approx f$ to allow active molecules to stimulate their specific receptors. There is some tolerance, since the probability of interaction by resonance corresponds to a peak that has some width.

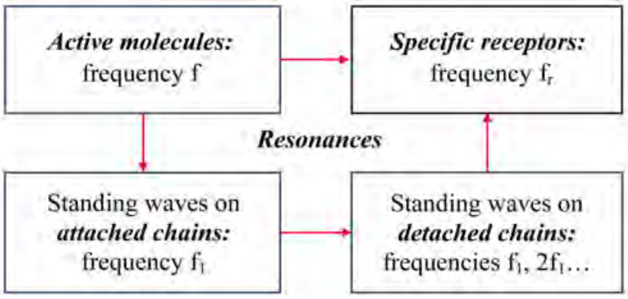


Figure 20. Water memory reveals a new type of molecular interactions.

The assumption that molecular interactions are only possible according to the "key and slot model" of chemical reactions is not correct. Biologically active molecules can also interact with their specific receptors by means of *oscillating electric fields and resonances*. This allow for a bypass, represented by the second line of **Figure 20**, since the oscillating electric field of active molecules can also create trimmed chains of WPs. As long as they are attached to their generators, they resonate at a frequency $f_1 \approx f$. This remains true when these chains are detached, but reactivation of their oscillations by thermal agitation in liquid water leads to a harmonic spectrum of possible frequencies ($f_s = sf_1$, where $s = 1, 2, 3, \cdots$). The fundamental frequency f_1 remains dominant, however. Standing waves on trimmed chains of WPs do produce an oscillating electric field of frequency $f_1 \approx f_r$ and can thus stimulate the same receptors.

The collective electric field, generated by all trimmed chains of WPs, has even the capacity to create more and more equally trimmed chains. Their number is increased and the oscillating electric field is amplified by an autocatalytic process. The possibility that molecular interactions can result from oscillating electric fields and resonance effects had been overlooked. The discovery of water memory did thus reveal the existence of a mechanism that is of fundamental importance and even very efficient.

The fourth erroneous assumption concerns homeopathy. Since the underlying mechanism was not understood, it was believed that its efficiency can only result from placebo effects. We wonder how they can be justified for animals and small children. Our purpose is not to defend homeopathy, but to restore truth, also in this regard. It is therefore instructive to examine the argumentation advanced by those who would like to eliminate homeopathy. The *Australian National Health and Medical Research Council* published in 2015 a study on "Evidence on the effectiveness of homeopathy" [41]. This report was sponsored by the Australian Government, but was not based on a scientific study of underlying physical and physiological process. It was merely a collection of 176 articles, supporting the claim that "there are no health conditions for which there is reliable evidence that homeopathy is effective."

These evaluations are essentially dependent on subjective appreciations. It was recognized that the general conclusion of their report was "based on all the evidence considered". Other evidence was discarded. The first report of 2012 had even been concealed [42], since it was not sufficiently selective in this regard. The published report stated that the aim of this study was "to assist people in making health care choices", but it was addressed to political deciders. The authors mentioned, indeed, that they wanted to "influence policy" and to get "increased funding for such research" ([41], pp. 6, 16 and 4). This report was widely publicized and is equivalent to lobbying.

3.5. The Standard Procedure for Extra-High Dilutions

It is even necessary to clarify the origin of homeopathy, which has often been misrepresented to denigrate it. The basic idea was due to *Samuel Hahnemann* (1755-1843). He was a regular medical doctor. After acquiring his diploma at the age of 24, he practiced during 5 years, but decided then to cease. He had realized, indeed, that it would have been better for some of his patients not to be treated according to the "art of healing" of his time. He was even horrified that he might "murder" suffering people, instead of helping them. This was an exceptionally honest attitude, justified by recognizing the cause of this horrible situation. Neither the chosen substances, nor the doses were determined in a rational way, although Paracelsus wrote already in 1543 that "only the right dose differentiates a poison from a remedy".

Hahnemann's linguistic gifts made it preferable for him to translate books and to search there for possible improvements of medical practice. In one of these books, it was claimed that the bark of a Peruvian tree was able to treat malaria. It is known today that the bark of "cinchona" trees contains *quinine.* Most efficient medicines were actually discovered by trial and error. It was already known in Antiquity, for instance, that leaves of willow trees can stop pain. A chemist discovered in 1853 that the active molecule is $C_9H_8O_4$, which became famous as *aspirin.* Even elephants, apes and other animals know how to cure or avoid ailments [43]. The textbook attributed the beneficial effects of cinchona powder to its taste, but Dr. Hahnemann could not believe this claim. Nevertheless, he decided to verify if there were any detectable effects and was amazed that it produced malaria-like symptoms.

By experimenting with other substances, he realized that medicines could be discovered in a more rational way, by adopting the "law of similars". This was merely an empirical rule, but such rules were often followed before understanding why they are valid. [Even Newton's law of gravity was expressed in terms of actions at a distance. It did account for observed phenomena, but the real cause is a gravitational field, which corresponds even to modifications of the metric of space and time.] Since Hahnemann tried to discover medications by means of tests, performed on healthy persons, he had to use the lowest possible doses. He adopted thus the method of successive dilutions. If the result was beneficial, such an EHD could also be administered to patients in a secure way. We recall that Hahnemann was a learned medical doctor and was thus able to verify if a preparation is helpful of not.

As an example, we mention *Apis mellifica*. Until recently, it was customary in medicine to use Latin, also for anatomy, to overcome language barriers. The European honey bee is called "*Apis mellifera*" and the main component of its venom is *mellitine.* This molecule has also anti-inflammatory properties and honey bees do even protect their larvae from infections by means of very efficient substances. Hahnemann presented his discovery already in 1796 in a German medical journal, by formulating the rule that "like cures like". Objections that are based on the finite divisibility matter are anachronistic. [The ancient concept of atoms had been reintroduced by Boyle in 1661 and elaborated by the chemist John Dalton in 1804, but the atomic theory was only accepted at about 1860, since the kinetic theory of gases did prove that Avogadro's hypothesis was correct. Nevertheless, Mendeleev did not yet dare to use the concept of atoms in 1869.]

Hahnemann could thus assume that even when a substance has been diluted many times, there remains something of this substance. In 1810, he presented a first collection of results and one year before his death, the 6th edition of his "Organon of the Rational Art of Healing" was ready for publication. It is easily available [44]. It should also be obvious that successive dilutions required always homogenization before the next step. Hahnemann did this by holding the vessel in his hand and stroking it with vigor on a semi-elastic surface, like leather. Modern chemists use mechanical vortexing to insure homogenization of mixtures. This is merely a simpler method. We will show in the following section that these "successions" do also have another effect. Although Hahnemann was only concerned with practical medicine, he discovered already the bypass of **Figure 20**. Neither he nor his detractors were aware of this fact. Even when Benveniste did empirically prove the biological efficiency of EHDs, it was categorically declared to be impossible.

3.6. The Quasi-Periodic Variations of Biological Efficiency

Sir John Maddox, long-term editor of Nature (1966-73 and 1980-95) accused Benveniste of self-delusion, although his article contained two figures, displaying results of measurements [1]. **Figure 21** reproduces one of them. It is undistorted

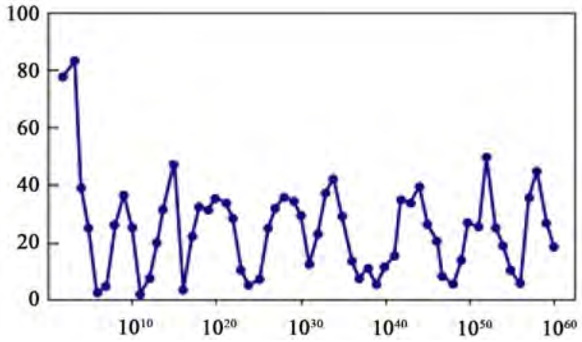


Figure 21. Measured quasi-periodic variations of the biological efficiency of EHDs, published by Benveniste *et al.* [1].

and was extracted from a publication in Japanese. We see 9 peaks. The first one is higher than the following ones. Since every peak did result from several measurements, the investigators should have realized that the quasi-periodic variations cannot result from "disregard of statistical principles" and "sampling errors". The investigators proclaimed even that Benveniste's experimental results were merely due to self-delusion [2]. What would result from the theory of water pearls? We begin with a description of the underlying processes in usual language by means of **Figure 22**. The first frame (a) shows some of the initially diluted active molecules, their charged parts (in red) and attached trimmed chains of WPs. We know that their length $L = \lambda/2$ and depends on the frequency f of the oscillating electric field of the active molecules. The second frame (b) illustrates the situation immediately after the first dilution and vigorous agitation. The concentration of active molecules has been reduced and agitation lasted long enough to detach all chains from the remaining ones, but some chains of WPs were broken.

The third frame (c) shows that after a relatively short time interval, the remaining active molecules had again formed trimmed chains of WPs. Broken detached chains did grow and new ones were generated by the global oscillating electric field. These chains have the same characteristic length L, allowing for standing waves at the frequency f_1 as well as harmonics $(2f_1, 3f_1, \cdots)$. Some chains may have reached the length $2L$. It allows for a mode of oscillation where $2L = \lambda$, which is equivalent to $L = \lambda/2$ and allows for the frequency f_1. There did also appear some "associated chains", resulting from mutual attraction of trimmed chains. **Figure 23** shows how two parallel chains of WPs will be attached to one another, but more than two chains are also possible. However, associated chains do not resonate at the same frequency or not at all.

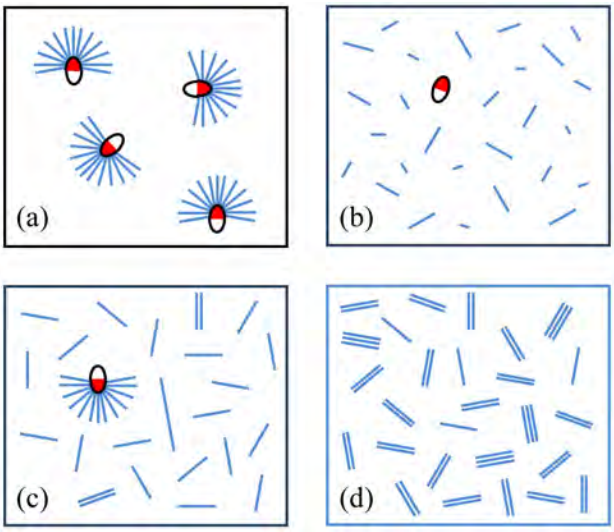


Figure 22. Decoding the standard procedure for preparing EHDs. (a) Active molecules and attached trimmed chains of WPs; (b) Detached and broken chains, immediately after dilution and agitation; (c) Reconstituted chains; (d) Excess of associated chains.

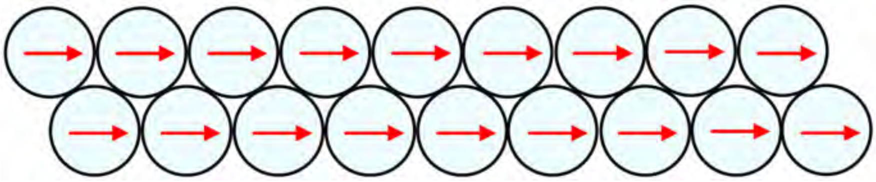


Figure 23. A simple associated chain of water pearls.

The fourth frame (d) of **Figure 22** represents a state where nearly all single chains of WPs got bound to one another. The biological efficiency of the resulting EHD is then much reduced, but associated chains can be broken by vigorous agitation. Single chains can then be multiplied again during successive dilutions, until there are so much of them that association will be favored. The global result is that the biological efficiency of EHDs has to vary in a quasi-periodic way. It is not perfectly periodic, since these processes allow for statistical fluctuations.

3.7. Kinetics of Water Memory

It is useful to express these ideas by means of equations, since they allow for logical deductions. Let X_o be the initial number of active molecules, dissolved in a given volume of twice distilled water. This concentration is reduced by successive dilutions, where the same fraction of the homogenized solution is eliminated at every step. Usually, this fraction is 9/10 or 99/100. It is replaced by pure water to get always the same volume. When successive dilutions follow one another at identical short time intervals Δt, the concentration of active molecules becomes a function $X(t)$ that decreases step-wise, since

$$X\left(t+\Delta t\right)=X\left(t\right)-\alpha\Delta tX\left(t\right)$$

The value of $\alpha\Delta t = 0.9$ or 0.99. For smalltime intervals, $X(t)$ can be treated as if it were a continuous function. It decreases then according to the equation:

$$\dot{X}=-\alpha X \quad \text{so that} \quad X\left(t\right)=X_o\exp\left(-\alpha t\right) \tag{11}$$

The exponential decrease does necessarily end up with $X(t) = 0$ when $t \gg 1/\alpha$, but this does not prove that the biological efficiency of EHDs has to vanish. Active molecules are able to generate trimmed chains of WPs with a probability g per unit time and they do generate more of them with a probability β par unit time. The concentration $Y(t)$ of trimmed chains of WPs increases thus according the equation:

$$\dot{Y}=gX+\beta Y \tag{12}$$

When the sequence of EHDs starts without previously formed trimmed chains, the initial value $Y(0) = 0$. Because of (11) and (12), we get then

$$Y\left(t\right)=\frac{A}{\alpha+\beta}\left[e^{\beta t}-e^{-\alpha t}\right] \quad \text{where} \quad A=gX_o \tag{13}$$

It appears that $Y(t) = At$ when $t \to 0$. The initial increase of $Y(t)$ is thus linear and very rapid when A is great. When the generation of new substitutes exceeds losses ($\beta > 0$), the concentration $Y(t)$ does eventually increase like $\exp(\beta t)$. This constantly accelerated increase would only stop when the whole amount of liquid water has been solidified. This might even apply to oceans and would be catastrophic, but is prevented by forming associated chains. Their concentration $Z(t)$ varies also, but

$$\dot{Y} = A\exp(-\alpha t) + \beta Y - \varepsilon ZY - \varepsilon' Y^2 \tag{14}$$

$$\dot{Z} = \varepsilon ZY + \varepsilon' Y^2 - \gamma_o Z \tag{15}$$

Equation (15) accounts for the fact that $Z(t)$ increases by combining already existing associated chains with single ones. The average rate ε is greater than for association of two single chains, because of more possibilities. (12) is replaced by (14), since associations imply that $Y(t)$ decreases by the same amount. However, every associated chain has also a probability γ_o per unit time to be destroyed by vigorous agitation. Figure 24 presents the results of numerical integration of (14) and (15). The unit of time is $\tau_o = 1/\alpha = 1$. We assumed that $A = 50$ and that $\beta = 0.5$, $\varepsilon = 0.1$, $\varepsilon' = 0$, $\gamma_o = 1.5$. The measured efficiency is proportional to $Y(t)$. This function is thus represented by a thicker line.

The red line describes the exponential decrease of the concentration $X(t)$ of active molecules and the thin blue line represents the variations of the concentration $Z(t)$ of associated chains. It was a "hidden variable" for Benveniste and his team. The first peak of $Y(t)$ is greater than the other ones, since $X(t)$ does still contribute to the generation of substitutes. The initial increase is linear and very rapid when $A = gX_o$ is great. This accounts for Figure 21. The following peaks were not always identical, but they had nearly equal heights. They are perfectly periodic in Figure 24, since we assumed constant parameters in (14) and (15). We assumed that $Y(0) = 0$ and $Z(0) = 1$, since $\varepsilon' = 0$ requires that one of the initial values $Y(0)$ or $Z(0)$ is different from zero to initiate the periodic variations. It should be noted that the peaks for $Y(t)$ are *dissymmetric.* The increase is progressive, but the decrease is precipitated, since single chains are more frequently associated with already existing ones when their concentration $Z(t)$ is great. This was not obvious in Figure 21, but has been demonstrated by Montagnier (Section 4.1).

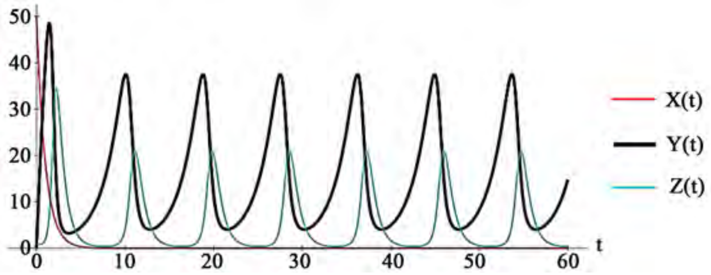


Figure 24. Decrease of the concentration $X(t)$ of active molecules and the periodic variations of the concentrations $Y(t)$ and $Z(t)$ of single and associated chains of WPs in the course of successive dilutions.

It is simply a matter of fact that if Maddox had tried to understand the experimental results, he would have discovered that similar variations were already known since 1910 for autocatalytic reactions and so-called "chemical clocks". Periodic variations attracted even more attention in the 1920th, since they were also observed for variations of the population density of predators and their prey. Predators proliferate when pray is abundant, but when the population of victims has been decimated, the predators have greater difficulties for their own survival and reproduction. Fewer predators allow the population of potential victims to grow again. This phenomenon was described by the famous *Lotka-Volterra equations,* which are identical to (14) and (15), when $A = 0$ and $\varepsilon' = 0$.

After developing the present theory, we found the book of Francis Beauvais [10]. He provided many examples of similar results obtained by Benveniste and his team and referred even to the Lotka-Volterra equations. Without the concept of WPs, he had to assume the formation of negative and positive molds. This hypothesis was only mentioned as an example, but proves that Maddox could have referred to the Lotka-Voltera equations, if he had tried to understand the unexpected experimental results.

Thomas Kuhn analyzed the process of scientific revolutions [45]. He distinguished different stages. When "anomalies" appear and cannot be explained by means of "normal science", the first reaction is to declare that they are impossible. Confirmation of the reality of these anomalies opens a period of "crisis". Various types of approaches are then tried and momentarily permitted. [For instance, Michelson's experiments of 1881 and 1887 did prove that the velocity of light with respect to the Earth is identical for any direction. This was not compatible with the concept of light waves propagating in ether with respect to the moving Earth. It was proposed at first that the ether could be entrained by the Earth, but Einstein realized that no luminipherous ether is needed. This hypothesis can be dropped, indeed, when we accept that results of space and time measurements depend on the chosen reference frame. Only the velocity c of light in vacuum has always the same value for any inertial frame.] This "paradigm shift" resulted thus from reexamining previous assumptions and correcting them. This is also necessary for water memory, because of the third assumption in Section 3.4.

3.8. Properties of the Limit Cycle and Aging

Consequences of (14) and (15) can also be expressed by displaying the variations of $Y(t)$ versus those of $Z(t)$. Any particular point (Y, Z) defines then the state of the system at some instant t and the evolution of this state is represented by a continuous line. The dark curve in **Figure 25** is equivalent to **Figure 24**. The green curve results from identical parameters, but starts at (0, 10), instead of (0, 1). Since the presence of more associated chains does immediately reduce the concentration of single chains, the first peak of **Figure 24** is reduced. However,

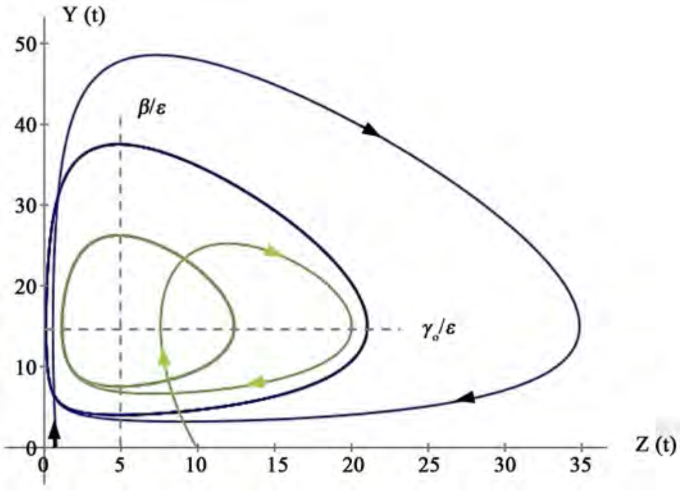


Figure 25. Parametric representation of the evolution of EHDs.

maximal and minimal values of $Y(t)$ are always reached when $Z = \beta/\varepsilon$. Those of $Z(t)$ require that $Y = \gamma_o/\varepsilon$. These facts result from (14) and (15) when $\dot{Y} = 0$, $\dot{Z} = 0$ and $\varepsilon' = 0$.

Figure 26 shows what would happen for $\varepsilon' = 0.0001$, when all other parameters have the same values as for **Figure 24**. We assumed again that $A = 50$, but the first peak of $Y(t)$ is greater, since $Z(t)$ was immediately increased by $\varepsilon' Y^2$. However, the amplitude of the periodic variations of $Y(t)$ is decreasing. The article of Benveniste [1] contained a second figure for a longer series of EHDs. It suggests a possible decrease of $Y(t)$, but ε' was surely smaller than 0.0001.

Equations (14) and (15) allow us also to answer two important questions concerning aging. Should the sequence of successive dilutions be stopped at an instant where the biological efficiency has a high value? Does the efficiency of homeopathic preparations not totally vanish after some time? To answer these questions, we solve Equations (14) and (15) after the instant $t = 0$, where the process of EHDs was stopped. We assume that all active molecules were already eliminated ($A = 0$), but without shaking, the values of β and γ_o are smaller. **Figure 27** shows the results of numerical integrations when β is reduced from 0.5 to 0.1 and γ_o from 1.5 to 1, while ε and ε' are not modified. The black curve results from $Y(0) = 30$ and $Z(0) = 1$, while the red one would be due to $Y(0) = 5$ and $Z(0) = 10$. We see that the initial conditions are irrelevant for the final result, although we assumed that $\varepsilon' = 0.0001$.

It follows indeed from (14) and (15) that $Y(t) \rightarrow \gamma_o/\varepsilon = 10$ and $Z(t) \rightarrow \beta/\varepsilon = 1$ for $\varepsilon' = 0$, when $\dot{Y} = 0$ and $\dot{Z} = 0$. The correction for small values of ε' can be obtained by introducing the lowest order approximation in the same equations. The reduction of the final values is negligible when $\varepsilon' = 0.0001$. It is remarkable that the alternative dominance of single and associated chains continues during the initial period of aging. We have also to stress the fact that preservation of the biological efficiency of EHDs requires that the system has not been perturbed. This can happen by heating and ultrasound, but also in another way.

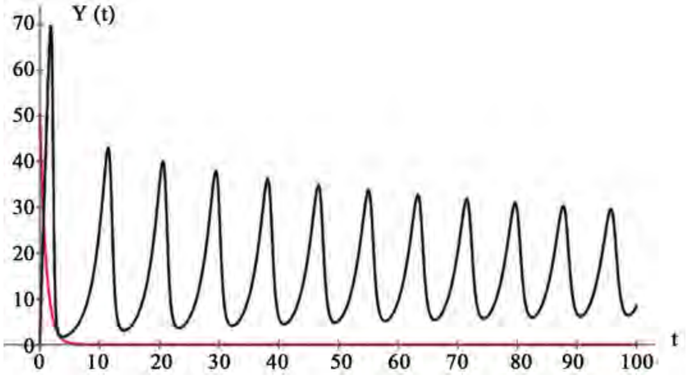


Figure 26. Possible variations of $Y(t)$ for many successive dilations.

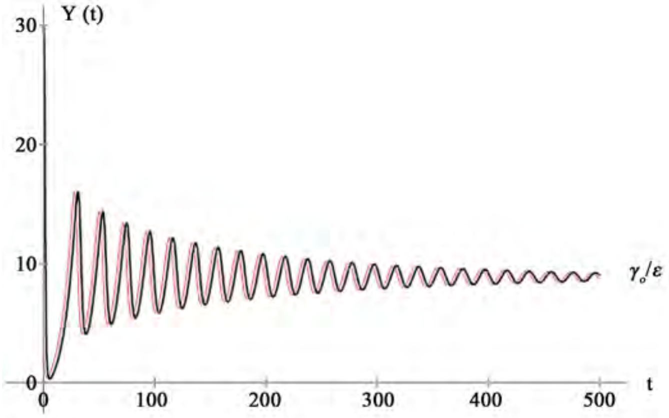


Figure 27. Effects of aging after stopping serial dilutions.

3.9. Crosstalk and Beat Phenomena

Official tests, performed in 1993, led to an unexpected fiasco (Section 5.2). It appeared, indeed, that the biological efficiency of EHDs was lost, while samples of pure water, needed for blind tests, turned out to be efficient. Benveniste was confronted to authorities, who concluded that the results did merely confirm their conviction that water memory is not a real and reproducible phenomenon. Since Benveniste knew that this was not true, because of numerous tests, he tried to understand the new observed facts. He realized that tubes with genuine EHDs of active molecules had been placed during some time near tubes that contained merely pure water. Benveniste thought therefore that the invisible structures, which are responsible for water memory, have to emit "signals", allowing them to transfer their biological efficiency to pure water. It was not clear why this information transfer was possible and why this could result in silencing authentic EHDs. However, Benveniste succeeded in proving, by means of purely empirical means, that the assumed signals do really exist.

He could detect them by merely putting a sample of some EHDs in a *coil.* This did yield an electric signal that could be amplified and applied to pure water. This involves physical processes that will be explained later on, but we know already that biologically active molecules stimulate their specific receptors by means of an oscillating *electric* field and resonance effects. Trimmed chains of WPs allow for standing waves and create an oscillating electric field that has the same effect inside EHDs. Normal interactions between active molecules and their receptors can also be bypassed (**Figure 20**).

We noted that resonance effects allow for some tolerance. It is thus sufficient that $f \approx f_1 \approx f_r$, where f_1 is determined by the length L of trimmed chains ($f_1 = v/2L$). However, it follows from Lo's measurements that $L = ND$, where $N \approx 20$. The fundamental frequency f_1 for standing waves will thus not always be precisely equal to the frequency f that is characteristic of the chosen type of active molecules. It can happen, for instance, that the chosen type of active molecules did *initially* create attached trimmed chains where $N = 20$ or $N = 21$. Autogeneration of equally trimmed chains of WPs during successive dilutions is governed by a collective electric field that oscillates then at one of these frequencies. EHDs are thus able to "breed" identically trimmed chains. However, the resulting "strains" can be slightly different for two EHDs of the same substance. This is irrelevant when $f \approx f_1 \approx f_r$. However, two EHDs with trimmed chains of nearly equal length produce electric fields of slightly different frequencies ($f_{\pm} = f_o \pm \Delta f$, where $\Delta f \ll f_o$). When these EHDs are contained in vessels that are put side by side, these fields are superposed. The resulting electric field does then oscillate at the average frequency f_o with a modulated amplitude. Indeed,

$$\cos\left(f_o - \Delta f\right)t + \cos\left(f_o + \Delta f\right)t = 2\cos\left(f_o t\right)\cos\left(\Delta f\right)t$$

$$\cos\left(f_o - \Delta f\right)t - \cos\left(f_o + \Delta f\right)t = 2\sin\left(f_o t\right)\sin\left(\Delta f\right)t$$

This *beat phenomenon* is well-known in acoustics. Even when the proximity of two slightly different EHDs was only temporary, their already strong electric fields generate together both types of trimmed chains. When this happened only during a short time, they will continue to produce both types of chains and therefore "mixed signals". Their biological efficiency will vary in a periodic way, but at a very slow pace. The relative phase of the superposed fields is also important, since the sum of two signals of equal amplitude will double their amplitude of oscillation. For the difference, the signals annihilate one another, but become strong again from time to time.

Momentary proximity of a genuine EHD with pure water can create there a small number of trimmed chains that resonate at one of the two possible frequencies. The sample of pure water becomes biologically active by "breeding" always the same strain of trimmed chains of Water pearls. Although Benveniste ignored the underlying mechanism, he had discovered that *information transfer* is possible. It proved the existence of signals and that they were responsible for water memory. Nevertheless, these signals had also extremely disconcerting effects. Beauvais used the term of "coherent discordances" to designate "wild transfers" and the fact that one operator did even "erase" the imprint [10] (p. 429, 450 and 453). These facts were enigmatic, but reproducible. For us, these parasitic effects do corroborate the theory of WPs.

3.10. Confirmations by Physico-Chemical Measurements

Vittorio Elia and his collaborators performed remarkable experiments with the "long term goal" of clarifying the problem of water memory [46]. They began in 1999 with measuring the *heat of mixing* when EHDs of active molecules (like As_4S_4) were dissolved in twice distilled water. It appeared that addition of EHDs of NaOH resulted in energy release. It was concluded that EHDs contain aggregates of water molecules that are able to create new ones. Elia insisted on the fact that these structures have the capacity of "self-organization". In terms of WPs, we can say that the added ions do also create trimmed chains of WPs. They are different, but bonds do always correspond to negative energy states. Mixing had thus to liberate more energy in the form of measurable heat.

The same team measured also the *electric conductivity* χ at 25˚C. Dissociation of NaOH produced ions that contributed to the measured electric conductivity. Ions liberated from the walls of the vessels did that also, but addition of EHDs of active molecules produced a significant *excess conductivity* χ^E. Since it was always proportional to the measured heat of mixing, both phenomena had a common cause [47]. The added active molecules increased the concentration of trimmed chains and thus also the electric conductivity inside molecular chains by the von Grotthuss mechanism.

Elia and his team made two other remarkable discoveries by measuring the excess conductivity χ^E for homeopathic dilutions of *Arnica Montana* during aging. The active substance is *helenalin*, ($C_{15}H_{18}O_4$, containing charged oxygen atoms). Its EHDs displayed wave-like variations of χ^E, but at an extremely slow pace [48]. In 4.4 years, there appeared only 4 peaks, which had nearly the same form, but their amplitude increased [49] [50]. We can now understand this fact in terms of beat phenomena. It appeared also that the excess conductivity χ^E was much greater for smaller volumes of EHDs. Initial conditions are usually irrelevant for aging (**Figure 27**), but $Y(t) \to \gamma_o/\varepsilon$. This value will thus be increased when the probability ε for creating associated chains is reduced by surface effects.

4. More Evidence of Ferroelectric Water Pearls

4.1. The Measured Frequency Spectrum

Benveniste discovered that EHDs of biologically active molecules produce "signals" that can be detected by means of a coil. The output was an electric tension that could easily be amplified and stored in analogical or digital form. The waveform was similar to that of noise, resulting from a superposition of simpler signals, randomly shifted with respect to one another. The predicted spectrum is a harmonic one:

$$f_s = sf_1 \quad \text{where} \quad s = 1, 2, 3, \cdots \ll N \tag{16}$$

This results from (9), which is an approximation of (10). Montagnier did publically show the results of Fourier analysis, nicely displayed on a computer screen [13]. The frequency spectrum did correspond to equidistant bell-shaped peaks.

Moreover, their heights varied in the course of successive dilutions. Montagnier's patent US2010323391 contains the spectrum obtained by Fast Fourier Analysis [51]. It yields clearly separated spectral lines for shorter extracts. For a particular type of active molecules, the reported values were

$$f_s = 1000, 2000, 3000, 4100, 5100 \text{ and } 5500 \text{ Hz}$$

The three lowest values do precisely correspond to (16). Since higher frequencies than f_1 result from molecular agitation in liquid water, they are excited with decreasing intensities. This implies greeter uncertainties. Important results of normal Fourier analysis were presented in an article [14], even for several substances and different stages of dilutions. They are especially interesting, like hieroglyphs, when we understand their meaning. For the first dilutions, there was only one great peak at the fundamental frequency f_1. We can attribute it to trimmed chains that are still attached to active molecules. Further dilutions led to the appearance of additional peaks, according to (16). Their average height increased during successive dilutions, but was maximal for dilutions D-9 to D-12. It did strongly decrease for D-13.

These facts agree with **Figure 24** and confirm the predicted dissymmetry of peaks of activity. For the purpose of these authors, it was not necessary to continue the sequence of dilutions. They wanted only to get EHDs of maximal efficiency. It could be reliably predicted by observing the height of the peaks at higher frequencies. It corresponds, indeed, to the first stage of successive dilutions where $Y(t)$ and the biological efficiency are maximal. It may be astonishing that the peaks for some harmonic frequencies f_s were depressed, but this due to interference effects, also for musical instruments [52].

4.2. Water "Balls" and Their Alignments

S.Y. Lo measured the diameter ($D = 15$ nm) of WPs. The second group of structures had variable sizes of about 300 nm. This yields the length L of chains of WPs. Montagnier and his collaborators tried to measure the size of the required information carriers of water memory by means of filters. This method suggested a size between 20 and 100 nm [15]. No one knew that the carriers of water memory are single chains of WPs and can thus be broken by filtering. If $L = ND$ were always shorter than 100 nm, these chains would contain at most $N = 7$ WPs. The resolution for distinguishing different types of active molecules would then be too small to be efficient. By the way, the measured values of $L = \lambda/2 \approx 300$ nm and $f_1 = v/\lambda \approx 1$ kHz, allow us to determine the velocity $v = 2L/f_1 \approx 0.6$ nm/s for any chain of WPs.

Lo found also by means of interference measurements a third group of supramolecular structures. Their size was much more variable, but of the order of 3000 nm [40]. The X-ray spectrum confirmed that they were only constituted of water molecules, even without silicon contaminations from the vessels. These structures could be created by solutions of different types of initiators, but required high concentrations of chains of WPs. Excitation by UV light produced fluorescence at 298 nm, which is absent for pure water and ice. These "water balls" survived after short heating at 80˚C, but were decomposed by vigorous agitation. They were then spontaneously reassembled after about 15 min. **Figure 28** reproduces, with kind permission of World Scientific Publishing Company, two pictures that were published the same year [53].

The left image proves that these balls tend to be aligned and the right image that they can be deformed. We propose therefore that water balls are constituted of chains of WPs, loosely bound to one another with global quasi-ferroelectric ordering. Water balls are thus dipolar, but contain also water molecules that can be expelled. They were said to be examples of "soft matter". Lo discovered also that when these balls are very numerous, they constitute extremely long alignments, visible by optical microscopy [54]. Spectacular images of these alignments were published in 2009 and reproduced in a book [55]. It gathers all articles of Lo and his collaborators concerning the discovery and empirical analysis of structured water. These images are also contained in the open-access article of Mae-Wan Ho [56].

It is very remarkable that these alignments are branching-off sideways, always at the same angle of 78˚. **Figure 29** shows that this astonishing fact can be explained. Since water balls contain chains of WPs, they have to be attached to one another at points, where all effective dipoles are oriented toward the center of neighboring water molecules. These dipoles are represented by red arrows, but parallel molecular chains are shifted with respect to one another by the distance *a*.

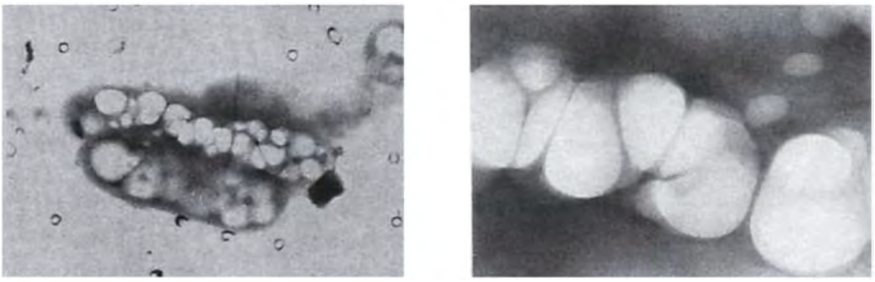


Figure 28. Transmission electron-microscopies, realized by Lo [51]. They show deformable "balls", constituted of chains of water pearls.

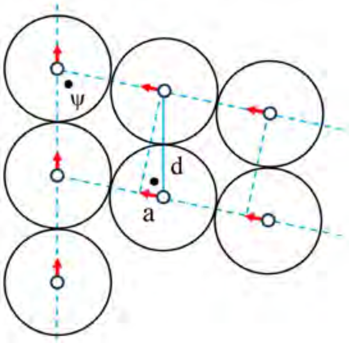


Figure 29. Chains of water molecules allow for strong bonds when they meet one another at the angle $\psi = 78˚$.

Since the centers of adjacent water molecules are separated by the distance d, molecular chains meet one another at an angle ψ, which appears in **Figure 29** at two different places, indicated by dots. It follows that $\cos\psi = a/d = 0.6/2.9$. Thus, $\psi = 78^\circ$. This result does strongly confirm the validity of the proposed theory.

4.3. Proposed Explanation of Water Bridges

What would happen if liquid water were subjected to very strong electric fields? This question has been raised long ago by the British lawyer *Willian George Armstrong,* who became a respected engineer, inventor and scientist. He had a powerful source of high electrostatic potential differences and used it to find out if they produce very intense electric currents in liquid water. He knew that these currents would result from motions of H^+ and OH^- ions, but he discovered surprising facts. He presented them in 1893 to a general audience as being entertaining, but mentioned that they might "be interesting to experts" [57]. Actually, he filled two wine glasses up to the rim with pure water and put them in contact with one another. He immersed a long cotton thread in one of them and introduced one wetted end of it in the other glass. He was amazed to see that a high potential difference between water in the adjacent glasses did transfer the whole cotton thread to the other glass. Why did this happen?

The spongy cotton thread had to be entrained by motions of charges. Since OH^- ions have a greater mass than H^+ ions, the thread should move towards the positive pole. This is what Armstrong observed, but the water level remained constant in both glasses. This was confirmed when Armstrong used a vessel where a thin tube allowed to see more precisely any variation of this level. He-mentioned that during a few seconds after the complete transfer of the cotton thread, a "rope of water" remained suspended between the two glasses. When *Elmar Fuchs* was studying physics in Austria, he heard about this phenomenon and reproduced it with a source that could sustain high currents at 15 kV, for instance.

The first results were published in 2007. He found that the cotton thread was not necessary to produce a "floating bridge" of liquid water [58]. It remained stable as long as the high electric field was applied. This phenomenon did raise a fundamental problem [59]: how is it possible that a strong electric field modifies the *interaction* between water molecules in the liquid state so much that it becomes able to resist gravity? The generation of the water bridge was itself a complex process. At first, the water surface became agitated in both vessels. Then, a very thin filament of liquid water established contact, but it lasted only a fraction of a second. Its diameter was then suddenly increased to yield a cylinder of circular section, indicating strong surface tension. It remained suspended in air like a catenary, even when its length was slowly increased up to 2.5 cm at 25 kV. Fuchs started then a systematic study of this strange phenomenon with various collaborators, to discover the cause of this "new state" of liquid water.

Since water bridges provided direct visual evidence, it would have been difficult to negate the reality of these observations, but the basic problem was the same as for water memory. What could be observed? In regard to electric conductivity, it was confirmed that the transport of charges is bidirectional [60]. Protons (H^+) and proton-holes (OH^-) are passing through the bridge. By measuring the complex impedance between 100 Hz and 10 MHz, it appeared even that the charges were moving like nearly-free conduction electrons in metals [61]. Raman scattering indicated also that vibrational modes in liquid water were not modified with respect to usual ones in bulk water [62]. Charge transport was due to the von Grotthuss mechanism, but we have to solve two problems. 1) Why is cohesion of water molecules in the liquid state enormously increased by applying a very intense electric field? 2) How can many H^+ and OH^- ions move in opposite direction, without hindering one another?

The first question has to be related to the higher density of liquid water near the surface of the bridge. This was proven by X-ray scattering for two-dimensional beams of submillimeter extension [63]. Since neutron scattering is sensible to hydrogen atoms, it yields complementary results [64]. The stability of water bridges has been attributed to higher surface tension [65], but that is not sufficient, since longitudinal stretching of the "liquid cable" requires greater bond-strength along this direction. We propose that the stability of water bridges is due to the formation of *very long chains of WPs.*

We can then understand the sequence of events. At first the water surface became agitated in both beakers, since growing chains were formed near the electrodes and then moving around, until contact was established between water in both vessels. At first, there were only few chains, constituting a capillary bridge, but once the way was open, more and more chains of WPs were rapidly formed by the very intense electric field. Statistical fluctuations of the traffic of protons and proton-holes along the thin bundle of chains of WPs led to repulsion. It became more efficient to push parallel chains towards the surface of the liquid bridge. The capillary bridge was, indeed, replaced by a thicker one to achieve a new equilibrium. Since WPs are constituted of more densely packed water molecules than in liquid water, chains of WPs, situated at the surface of water bridges should there produce a greater density. This explains also the observed birefringence for linearly polarized light. Moreover, higher potential differences and greater electric currents require more chains of WPs at the surface of the liquid bridge. This should increase the diameter of the bridge and does agree with observations [66].

An increase of the applied potential difference allows also for a greater length of the catenary. This fact is due to the polarization of water molecules, as shown in **Figure 30**. The left part represents a molecular chain like that of **Figure 10** for a moderate electric field, while the right part indicates that a very intense electric field **E** does strongly polarize the water molecules. The usual angle of 106˚ between OH bonds is reduced and the distance between proton pairs and the negative cores of oxygen atoms is also reduced by deformations of the usual, nearly

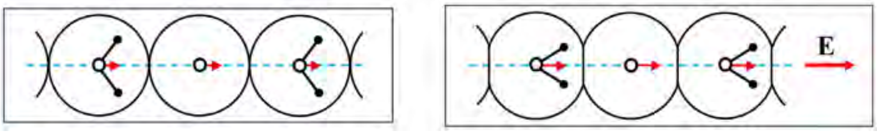


Figure 30. Aligned water molecules get more tightly bound by a strong electric field **E**.

spherical shape. Very strong polarization of water molecules would thus increase the mutual attraction of neighboring water molecules. This is also true when the bonds are due to intramolecular exchange effects (**Figure 6**). Electric conduction by the von Grotthuss mechanism would thus be facilitated, but regulation of opposite traffic of protons and proton holes is imperative for very intense electric currents passing though the water bridge.

Some observations indicated that this might happen in a helicoidally coordinated way [65]. This makes sense, when we adopt the model of **Figure 31**. Groups of positive and negative charges are there represented by red and blue dots. They are distributed in a plane, to constitute a lattice like that of ionic crystals, but this plane rotates around the axis, while the charges are advancing along different lanes, represented by red and blue lines. [This double helix is similar to Da Vinci's famous stair case of the chateau of Chambord in France. It contains two flights, allowing for simultaneous upward and downward traffic.]

4.4. Lack of Evidence for Coherence Domains

Emilio Del Giudice et al. proposed in 1988 a bold hypothesis concerning properties of bulk liquid water at very small scales [67]. Basically, they assumed that inside some water molecules that belong to small spherical domains, electrons can remain in an exited state. Since the ionization energy of water molecules is 12.6 eV, they considered excitation to an energy level at 12.06 eV. To keep the electrons in this weakly bound state, they assumed constant emission and reabsorption of *virtual photons,* trapped inside the small spherical volume. This required that electrons should be accumulated at the surface of the sphere, while positive charges had to be expelled and would be homogeneously distributed in the surrounding liquid water. The accumulated electrons should constitute a barrier that reflects virtual photons. It was also necessary to postulate that the ensemble of excited electrons inside such a sphere are oscillating in phase with the EM field of all virtual photons. Because of the selected excitation energy, the wavelength would be close to 100 nm and determine the size of these *coherence domains* (CDs). It was estimated that they contain about 1000 water molecules, since only a fraction of the internal electrons would be excited.

Although this theory was initially conceived for bulk water [68], it was assumed later-on that it should even account for exclusion zones, water bridges and water memory. Because of the prestige of quantum field theory and the need of a rational justification of the mysterious phenomenon of water memory, the theory of CDs had increasing impact. However, we have also to mention what happened in 1993. It had already been decided that Benveniste's Unit 200 would

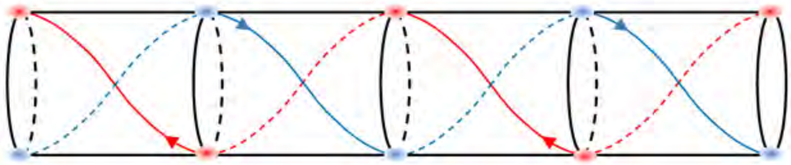


Figure 31. Opposite motions of bunched charges along a double helix.

be closed, but he found a more objective method for detecting the biological efficiency of EHDs. He was thus allowed to prove the reality of water memory by means of blind experimentation, subjected to rigorous control of *Georges Charpak*. He was the 1992 Nobel Prize laureate in physics, since he invented a new type of particle detectors for CERN. Benveniste tried to convince him, by referring to the article of Del Giudice *et al.* [67].

Since this article concerned condensed matter physics, Charpak asked the opinion of *Pierre-Gilles de Gennes*, also Nobel Prize winner in physics. He answered that this theory is "worth nothing", since it is based on "false hypotheses" ([6], pp. 154, 168). Beauvais mentions that de Gennes had asked *Philip Nozières* ([13], pp. 395, 398 and 402). He was the top theoretical physicist in France for new properties of condensed matter. Since such a harsh judgment might have been influenced by controversies, we have to know why the concept of CDs has been proposed. The theoretical physicist *Giuliano Preparata* (1942-2000) made important contributions to elementary particle physics, before becoming interested in applying quantum electrodynamics (QED) to condensed matter physics [68]. He viewed the existence of CDs as being a general property of pure liquid water in the sense of a "two fluid picture". He stated even that "the concept of H-bonds is merely phenomenological", since liquid water should rather be viewed as a "condensed vapor" than "molten ice" ([68], pp. 38, 196, 197). Liquid water would locally be condensed, before becoming ice.

Preparata was coauthor of the official article on CDs [67], presenting the same idea in Physical Review Letters. The primary author, *Emilio Del Giudice* (1940-2014), was one of the pioneers of string theories and explained why he got interested in CDs [69]. Since his youth, he did feel that the whole universe constitutes a profoundly integrated entity and studying physics, he developed a sense of wonder in regard to quantum mechanics (QM). He got especially interested in "entanglement". Since this concept provided his motivation for considering CDs, we have to clarify its origin and physical meaning.

Albert Einstein had contributed to the development of quantum mechanics, but in 1927 he learned about new ideas and discussed them with Niels Bohr at the Solvay conference in Brussels. He perceived very keenly that this theory attributed peculiar properties to measurements. He developed this idea in the famous EPR article of 1935, by means of a thought experiment. We describe it in equivalent terms, by considering two particles that have a property that can be precisely measured, but allows only for two possible values: +1 and −1. Since QM accounts for limited knowledge, we can define a state where it is only known that the values ±1 are equally probable. It is then sufficient to determine by means of a new measurement that the value is +1 for one particle, to be instantly sure that it is −1 for the other particle. It does not matter how far these particles are separated from one another at that instant.

In classical physics, that would require the existence of a physical link. Einstein asked therefore: are such "spooky actions" at a distance physically possible or not? He did only raise the problem, while Schrödinger had developed a theory where causal relations were preserved for the propagation of wave functions in space and time. He insisted on the need of causal relations and did not like the idea of "quantum jumps". They do not allow for further analysis. Since Einstein had described in a vivid way that some measurements seem to imply universal connections, he coined the word of "entanglement" to account for them. They would require that instantaneous information transfer is possible at any distance, although this is excluded by the theory of relativity. Actually, it is sufficient to accept that QM is a theory of *possible knowledge,* limited by universal restrictions that Nature imposes on some measurements.

Classical physics postulated that totally precise, simultaneous knowledge of positions and velocities is possible. It implied continuity and strict causality, but this assumption has to be corrected. The basic paradigm of QM is that the constant h imposes irreducible uncertainties. The concept of "virtual photons" results from the fact that during short time intervals Δt, the energy of a system can only be known with a minimal uncertainty ($\Delta E \approx h/\Delta t$). Since the theory of relativity imposes that the energy (mc^2) of particles of given mass is finite, the number of identical particles cannot be precisely known during short time intervals. This applies also to photons and leads to some observable effects.

Del Giudice was convinced that virtual photons are even able to establish a link between physical systems when they are separated by arbitrarily great distances. He expressed this idea [70] by stating that *telepathic communications* are possible, but he knew that physical theories are only justified by their agreement with observations. He tried thus to show that the concept of CDs accounts for unexplained phenomena, like water memory, exclusion zones and the stability of floating water bridges. This idea was diligently propagated, but does it really account for water memory?

Bellavite and other health specialists at Italian universities reviewed in 2013 the status of research concerning EHDs [71]. They concluded that "evidence strongly supports the notion that the structuring of water and its solutes plays a fundamental role", but the mechanism is not yet understood. Two hypotheses had been proposed: hydrogen bonded clathrates and coherence domains. After examining the evidence, they concluded that "none is fully convincing". The concept of ferroelectric water pearls was not yet known. It explains essential properties of EHDs, like the observed quasi-periodic variations of their biological efficiency, the measured frequency spectrum of signals created by local structures and even the angle of 78˚ between alignments of water balls. In his regard, we have to mention that *Mae-Wan Ho* tried to explain alignments of water balls [56], by assuming that expelled protons are not homogeneously distributed in liquid water. They could be more concentrated between the charged CDs, but that is not sufficient to explain the observed angle of 78˚ for chains of water balls.

Benveniste was sure that water memory is real, because of the often verified quasi-periodicity of the biological efficiency of EHDs and detection of signals. It was thus obvious that a physical explanation had to be possible. Since CDs seemed to offer an explanation, Montagnier accepted the help of scientists who advocated this approach. They stated even that water memory involves the "gauge theory paradigm of quantum fields" and "the framework of spontaneously broken gauge symmetry theories" [72]. These words are familiar to those who know the jargon of quantum electrodynamics (QED). They are impressive, but are they necessary? We recall that Thomas Kuhn distinguished three periods in the process of scientific revolutions [45]. At first, the observed anomaly is declared to be impossible, since it does not agree with conventional knowledge. When it is confirmed and cannot simply be denied anymore, it comes to a "crisis", where any hypothetical explanation may be proposed.

The theoretical chemist Tamar Yinnon published a series of articles, where the concept of CDs was elaborated, by postulating the existence of various types of these structures. He presented them in 2015 as being "predictions of QED", but we found only a catalogue of assumed structures [73]. The original CDs were renamed CD_{elec}. Their diameter would thus be about 100 nm. Greater structures, called CD_{plasma}, were assumed to contain some molecules of the solvated substance, but also more water molecules. Their effective dipoles would be oriented by the electrically charged active molecules or ions inside these domains. The resulting hydration spheres were said to be subjected to (monopolar) plasma oscillations with overall coherence. The size of these domains would be of the order of 1000 nm = 1 μm. A third type of CDs should reach sizes of 10 - 100 μm. They were called CD_{rot} and assumed to contain only water molecules, but all of them would have nearly parallel effective dipoles. They would be large elongated ferroelectric particles. Moreover, agglomerated CD_{plasma} and CD_{rot} entities constitute "supra-domains".

4.5. Forced Oscillations of Water Molecules in EHDs

Instead of commenting these respectable attempts to find an explanation of water memory, we continue to collect experimental results and to test the validity of the concept of WPs. *Adriana de Miranda* measured, for instance, the dielectric response of water molecules that interact with supramolecular structures [74]. She used 14 lots of LiCl, prepared by pharmaceutical laboratories and determined the spectral distribution of the real and imaginary parts of the impedance in the frequency range of 1 kHz - 13 MHz. She defined these variables by means of an equivalent circuit and compared the results to those of the measured dielectric response of EHDs of H_2O molecules in pure water, prepared in the same way. To understand the underlying physical processes, we consider the response of the effective dipoles of water molecules to an oscillating electric field at various frequencies. The center of these molecules remains practically motionless, while the positive tip of the effective dipole is displaced by a small distance $u(t)$ along the direction of the applied electric field. When n is the density of water molecules, the instantaneous polarization density is $P(t) = n2qu(t)$, where the displacement u is subjected to the equation of motion

$$\dot{u} = -u/\tau + (q/m)E(t)$$

Indeed, $2q$ is the charge and $2m$ the effective mass of the pair of protons. When the applied electric field $E(t)$ is suddenly extinguished at the instant $t = 0$, $u(t) = u(0)\exp(-t/\tau)$. The value of the *relaxation time* τ depends on all possible interactions in bulk water. For an electric field that oscillates with some constant amplitude at a given (angular) frequency ω, we get $E(t) = 2E_\omega \cos(\omega t) = E_\omega \exp(-i\omega t) + c.c$ and the instantaneous polarization density is $P(t) = \beta(\omega)E_\omega \exp(-i\omega t) + c.c$. The complex conjugate ($c.c$) requires merely that ω is replaced by $-\omega$. It follows that

$$\beta(\omega) = \frac{n(2q)^2/2m}{\gamma - i\omega} = \frac{\beta(0)}{1 - i\omega\tau} = \frac{\beta(0)(1 + i\omega\tau)}{1 + (\omega\tau)^2} = \beta_1(\omega) + i\beta_2(\omega)$$

The function $\beta_1(\omega)$ specifies the average orientation of effective dipoles. At low frequencies ($\omega\tau \ll 1$), they are oriented along the direction of the applied electric field. The polarization drops quite suddenly when $\omega\tau \approx 1$ and is reduced to zero when $\omega\tau \gg 1$. The imposed rhythm is too fast to allow the molecules to follow, because of friction. The imaginary part $\beta_2(\omega)$ varies like $x/(1 + x^2)$, when $x = \omega\tau$. This yields a peak that is centered on $x = 1$. The function $\beta_2(\omega)$ describes the *energy-loss* of water molecules, because of friction. It is maximal when $\omega = 1/\tau$. Since increased friction leads to lower values τ, a higher frequency is then required to achieve great energy losses. In QM, higher frequencies correspond to higher energies and in QED even quasi-static electric forces are due to exchanges of virtual photons. Miranda found that the values of relaxation times were situated between 40 and 100 kHz for EHDs of LiCl and H_2O molecules in pure water. This means for us that friction resulted from the creation of trimmed chains of WPs that had different lengths. However, the height of the peak for the energy-loss function varied in the course of successive dilutions of LiCl in such a way that it was maximal for D-9. Since friction is proportional to the concentration $Y(t)$ of single trimmed chains of WPs, this agrees with our theoretical predictions (**Figure 24**).

4.6. Discovery of Highly Remarkable Resonances

Since we expected that stationary waves on trimmed chains of WPs can be excited by an electric field, we did search relevant data and found the results of the Indian electrical engineer *Chitta Ranjan Mahata* [75] [76]. He had witnessed the healing power of homeopathy and wanted to understand how this is possible. Studying the literature [77], he discovered the proposition of *Barnard* [78]. He declared already in 1965 that homeopathy is in need of a "surer foundation" than to refer to positive results. This should be possible, since homeopathy requires that EHDs create substitutes of the initially dissolved active molecules. Since they should merely be constituted of water molecules, Barnard thought that they might "join up to form long molecular chains". This idea was probably suggested by the concept of polywater, which was in vogue at that time. Barnard proposed that these molecular chains may be wrapped around active molecules and that their length would thus provide the required specificity.

We were surprised that some important ideas had already been formulated more than 20 years before Benveniste's discovery. Specificity was even related to the length of local structures, composed of water molecules. Barnard was aware of the hypothetical nature of this proposition, but insisted that it indicates "the kind of experimental research in physics and chemistry needed now to establish the truth of homeopathy." Mahata realized that he could test this hypothesis, since ordered molecular groups in liquid water should lead to resonance effects.

Before describing and explaining his results, we have to mention that Barnard and Stevenson provided more details in another article [79]. They reviewed what was known about properties of EHDs, prepared for therapeutic purposes. It had already been established that their capacity of healing is subjected to "rhythmic variations" in the course of successive dilutions. It was also mentioned that biologically active molecules contain electric charges and that Alphonse Gay had developed in the early 1950th a method for measuring the dielectric response of EHDs [80]. It was complicated, but sufficient to identify EHDs of NaCl after 27 centesimal dilutions. [We thank Dr. van Wassenhoven for his help to get access to this paper.]

Mahata was not aware of these measurements and developed with his collaborators a much more efficient technique. In 2007, it was ready [81] and in 2010, they presented results for the spectral distribution of the real and imaginary parts of the dielectric constant [75] in a large frequency domain (100 kHz - 50 MHz). They found typical features of resonances for preparations of *Cuprum met,* but also for pure water. In 2013, they provided more results [76] and did interpret them in terms of Barnard's model. There were resonances, but we noted four perplexing anomalies:

1) These resonance frequencies were always very high: about 25 MHz instead of the expected ones at about 1 kHz.

2) There was only one resonance frequency, without harmonics.

3) A resonance was even observed for pure water.

4) The energy-loss functions were not symmetric, as required for usual resonances.

This did not correspond to expectations for single chains of WPs, but we realized that the third anomaly could be explained, since pure water contains H^+ and OH^- ions. Their electric fields might be sufficient to produce at least isolated water pearls. The observed high value of the resonance frequency would then require a very strong restoring force. For single WPs, it could only be due to their surface charges. They are represented by red and blue rims in **Figure 31**. For low resonance frequencies (~1 kHz), it was licit to assume that the oriented water molecules remain oriented towards the surface of the WP. However, at high frequencies (~25 MHz), the solid substrate is moving too fast to allow them to be reoriented so rapidly. Miranda's measurements indicated, indeed, that reorientation of water molecules occurs at lower frequencies. Because of the proximity and spatial distribution of the internal and external charges, this would then lead to a very strong restoring force.

We test the validity of this hypothesis, by comparing its logical consequences to Mahata's experimental results. The component E(t) of the applied electric field, which is normal to the symmetry axis of the WP sets the positive and negative poles of the WP in forced oscillation. The equation of motion for small displacements $u_o(t)$ of these poles is

$$\ddot{u}_o + \Omega^2 u_o + \gamma \dot{u}_o = C_\omega \exp(-i\omega t) \tag{17}$$

The (angular) resonance frequency Ω is determined by the strong restoring force and the effective inertial mass of the poles. Energy losses by friction are characterized by γ, while C_ω is proportional to the amplitude of the electric field E(t). Thus, $u_o(t) = B(\omega) E_\omega \exp(-i\omega t)$ for constantly forced oscillations. The polarizability β(ω) of pure water that contains a given concentration of isolated WPs is proportional to B(ω). When we normalize β(ω) to get always the same static polarizability β(0) = 1, it follows that

$$\beta(\omega) = \frac{\Omega^2}{\Omega^2 - \omega^2 - i\omega\gamma} \approx \frac{\Omega^2}{\Omega^2 - \omega^2} + \frac{i\Omega\Gamma/2}{\Omega^2 - \omega^2 - \Gamma^2} \tag{18}$$

We simplified the real part $\beta_1(\omega)$ to insist on the fact that without friction, this function would diverge when ω = Ω. At low frequencie (ω ≤ Ω), the average orientation of an ensemble of isolated WPs is identical to that of the applied electric field. It is opposite when ω ≥ Ω, but WPs cannot follow the applied field when $\omega \gg \Omega$. The function $\beta_2(\omega)$ describes *energy losses.* They are maximal when $\omega \approx \Omega$. The peak is symmetric and we can set γ/2 = Γ. The height of the peak is then Ω/2Γ and its width at half height is equal to Γ. **Figure 32** shows (in black) the frequency dependence of β_1 and (in red) the spectral distribution of the energy-loss function β_2 when the resonance frequency f = 25 MHz and Γ = 5 MHz.

These results had to be expected for a normal resonance, but Mahata's experimental results for β_2 did correspond to curves like the blue one. The height of the observed peak has been adjusted in **Figure 33** to coincide with the height of the normal peak and $\beta_{2\exp} \to 0$ when $\omega \gg \Omega$. Measured energy-loss functions are thus *asymmetric.* This feature did also appear for EHDs of various biologically active substances. Only the values of Ω and Γ, as well as the mysterious "dip" of the energy loss functions were slightly different. All these facts will be explained later on (in Section 5.4). Before we do that, we continue the equally necessary search of more evidence.

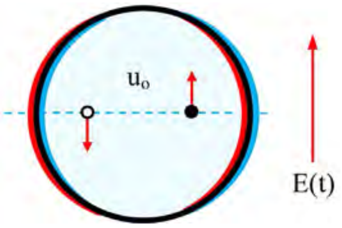


Figure 32. Rapid oscillatory rotations of a single water peal, excited by an electric field.

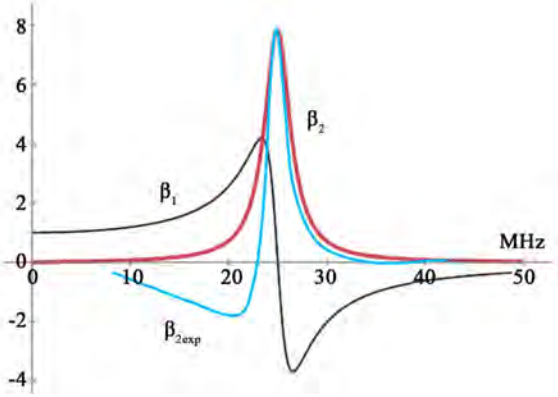


Figure 33. The spectral distributions of the real and the imaginary parts of the polarizability β, predicted for oscillators that resonate at 25 MHz. The blue curve represents typical results for Mahata's resonances.

4.7. Complementary Experimental Evidence

The Swiss biochemist *Louis Rey* did prove in 2003 that EHDs of NaCl and LiCl in ultrapure water produce local structures that are preserved after freezing [82]. He kept the EHDs during 24 h at −20˚C and reduced then the temperature to 77 K. When he exposed the transparent disk to a high dose of ionizing radiation, most excited electrons did immediately return to their normal state in ice, but some of them were trapped in local defects. To leave enough time for diffusion of electrons, the ice remained during one week at 77 K. After reaching equilibrium, the sample was warmed up at constant rate. The trapped electrons were thus progressively raised to states of higher energy, allowing for light emission. This *thermoluminscence* is a standard method for examining defects in insulators, since it determines the excitation energies that are required to liberate electrons from various traps.

There appeared a glow for $T \approx 120$ K and a more intense one at $T \approx 166$ K, especially when NaCl had been dissolved in heavy water (D_2O). Since electron traps were different for NaCl and LiCl, at least some of them were due to Na^+ and Li^+ ions. Rey concluded that these results prove "without any ambiguity" that liquid water has been structured in a lasting way. Since the glows appeared only when serial dilutions were followed by vigorous shaking, Rey thought that it might produce *nanobubbles*, attracted by ions [83]. The basic facts did not depend on this interpretation and were brilliantly confirmed by van Wijk *et al.* [84]. We can conclude that EHDs did produce trimmed chains of WPs, where the molecules are more closely packed and so strongly bound to one another that the ferroelectric crystallites do even subsist in normal ice (**Figure 4**).

Demangeat *et al.* studied the effects of EHDs by means of *nuclear magnetic resonance* [85]. Protons have indeed a magnetic moment that tends to be parallel or antiparallel to the direction of a constant magnetic field $\boldsymbol{B}_o$. The energy difference is proportional to the magnitude of this field and the transition can be stimulated by EM radiation of adequate frequency. When the excitation ceases, the system returns to its ground state, but there are two different relaxation times, T_1 and T_2, for components of the magnetic moment of protons along the direction the magnetic field $\boldsymbol{B}_o$ and perpendicular ones. The ratio T_1/T_2 depends on their environment. For EHDs, the results of measurements did prove that "water is a self-organizing system" [78]. *Van Wassenhoven et al.* presented more detailed results of NMR studies, confirming this conclusion [86] and providing additional information.

Other very important facts were discovered in Russia. They did prove that water molecules can also constitute stable structures without having to apply the standard procedure to get EHDs. *Konovalov and Ryzhkina* presented in 2014 a review [87], where they stated that the following conditions have to be satisfied: 1) only one type of molecules should be dissolved in pure liquid water at extremely low concentration, 2) these molecules have to carry an electric charge and 3) be subjected to low-frequency EM fields. This was verified for different types of molecules and the resulting structures were also observed by Atomic Force Microscopy (AFM). The authors insisted that these entities do not result from the formation of nanobubbles, but are associations of water molecules. This is essential to justify the concept of WPs, although the hypothesis of nanobubbles is closer to conventional ideas.

Burkin and his collaborators [88] performed light scattering measurements for very low concentrations of diluted molecules and interpreted their results in terms of bubble clusters around ions. This hypothesis leads to inconsistencies [89]. The Indian chemical engineer *Chikramane* and his collaborators used homeopathic drugs, prepared by means of EHDs of various metals (Au, Ag, Cu, Sn, Zn and Pt). They did yield local structures of sizes that ranged from 214 to 325 nm. They could thus be trimmed chains of WPs of length $L \approx 300$ nm, but these authors claimed that x-ray analysis demonstrated that they do contain particles of the initially dissolved metals. They assumed therefore that minute gas bubbles were formed during agitation and stabilized by attachment to small metal particles. Moreover, they thought that these entities rise to the surface and constitute there a thin layer that is reconstituted during subsequent dilutions. This would raise unanswered questions. NMR and other measurements of Van Wassenhoven for copper particles did not confirm this hypothesis [86]. Maybe, some procedures for preparing EHDs allow that charged metal particles get attached to WPs, because of their surface charges.

Demangeat thought also that vigorous agitation could produce nanobubbles. He proposed even that they might account for water memory by creating a "stereospecific shell" around active molecules [90]. How could these molds be detached, remain stable and allow for replication? Nevertheless, the concept of nanobubbles was attractive in the context of the "crisis", which resulted from the unsolved puzzle of water memory. It required a specific experimental investigation. The Slovenians *Sedlák and Rak* combined light scattering with rapid centrifugation and established that biologically active molecules at very low concentration produce real material entities [91]. Their measured sizes ranged from100 nm to 400 nm for various substances. It was explicitly noted that the mechanism of formation of these supramolecular structures and their nature were not yet known. However, the exclusion of nanobubbles and the reported sizes are compatible with trimmed chains of WPs.

Elia *et al.* produced supramolecular aggregates in pure water by repeated contact with a polymer, called Nafion [92]. It is strongly hydrophilic, because of its electrically charged sulfonate group (-SO_3H). It produced structures in liquid water, subsisting after freeze-drying. In solution, they had a high electric conductivity, attributed to proton hopping. They exhibited UV absorption at 270 nm and modified IR absorption, associated with the OH stretching mode of vibration. Solid residues after evaporation displayed clustered particles of about 40 to 400 nm. It appeared that these aggregates of water molecules produce *circular dichroism* [93]. This means that plane polarized light, decomposed in left and right circular polarized light, does not lead to equal absorption of these components. Dichroism indicates the presence of left-handed and right-handed structures, which can be tripoles in ferroelectric crystallites.

5. New Types of Molecular Interactions

5.1. The Sense of Smell

The most important consequence of water memory and its elucidation is that molecules do not only interact with one another by direct contact. This discovery belongs to a trend that began with trying to understand the sense of vision. John Dalton had described in 1794 his color blindness. Actually, he was unable to distinguish green from red. The British physician and physicist *Thomas Young* was intrigued by this anomaly, which led him to raise again the basic question: what is light? Newton had discovered that light of different colors can be separated by refraction. This could be explained by assuming that light is composed of particles, moving at constant velocity in any homogeneous, transparent medium. Why is it constant and depends on the medium remained mysterious, but refraction would then simply result from acceleration or deceleration at the interface.

However, Young discovered in 1801 that when light passes through two very narrow and close holes or slits, there appear dark fringes. He explained these results in terms of superposed "undulations", similar to those that can be observed on water surfaces. These interference effects led to the concept of *light waves*. Colors were then determined by their wavelength. They can be measured by means of gratings, but our eyes do not perform this kind of spectroscopy. Young realized that color vision requires only three types of receptors in our retina, mainly sensible to red, green and blue. Dalton had no green sensible receptors. These receptors had to absorb energy. [During about one century, it was assumed that light corresponds to waves. Actually, it is composed of photons, which are particles that carry energy and momentum, defined in terms of frequencies and wavelength. Photons behave even according to laws that are valid for waves. This synthesis transcends the idea of a dual nature of light.]

The English physiologist *William Ogle* tried to understand the sense of smell. Since anomalies could provide a clue, he collected and analyzed cases of "anosmia", *i.e.* partial or total loss of the sense of smell. He concluded in 1870 that odors are not perceived by means of chemical processes. It requires receptors that detect waves [94]. Ogle proposed that they should correspond to "undulations of the ether". [The concept of EM waves emerged only in 1873, through the development of Maxwell's theory. He gathered the laws of electricity and magnetisms, established for static or slowly varying electric and magnetic fields. He found that they are inconsistent when they are extrapolated to high frequencies, but harmony could be restored by adding a new physical law. It implied that electric and magnetic fields can constantly be transformed into one another, without needing the presence of electric charges. The theoretically predicted existence of electromagnetic waves was experimentally confirmed in 1887 by *Heinrich Hertz.*]

Because of QM, it became also clear why molecules can emit photons of infrared light and that this type of spectroscopy allows us to distinguish different types of molecules from one another. *Malcolm Dyson* proposed therefore in 1938 that the sense of smell is due to receptors that detect vibrational frequencies of molecules by energy absorption [95]. How is that possible without complicated instruments? This problem has been solved in 1996 by the biophysicist *Luca Turin.* Being a specialist of perfumes, he wanted to understand how odorants can be perceived and differentiated. Since olfactory receptors of insects, animals and humans are very small, he was looking for very compact detectors. Thus, he explored the domain of solid state electronics. Eventually, he found that this can be achieved by means of inelastic electron tunneling [96]. He told the story of this discovery in a very interesting book [97].

The Japanese *Leo Esaki* had invented in 1957 an efficient diode. It was based on properties of two n or p type semiconductors, separated by a very thin gap. It blocks the passage of electrons or electron holes, when the conduction band on one side meets a forbidden band on the other side. However, these bands can be shifted with respect to one another by applying a potential difference. Charge carriers can then pass through the intermediate potential barrier by wave-mechanical tunneling. The Norwegian *Ivar Giaever* applied this method to prove in 1960 that the BCS theory of low temperature superconductivity is correct. It had assumed that electrons can constitute bound pairs and predicted the existence of a forbidden band for possible energies. Giaever shared the Nobel Prize with Esaki for demonstrating that this is true.

Somewhat later, physicists realized that it is not necessary to apply a potential difference when the passage of electrons through a very thin gap between two semiconductors is not possible. It is sufficient that electrons of higher energy lose some kinetic energy inside the gap by collisions with other particles. Turin understood that this method allows us to distinguish molecules from one another, since this is equivalent to determining the energy required to excite vibrations inside these molecules. This is a simplified version of infrared spectroscopy, but requires specific receptors for different odors. That is a matter of genetics, as for color vision. Our color-space is usually three-dimensional, while the odor-space is multidimensional. Honeybees have 174 types of receptors and ants have even about 400 different ones [98]. This fact led to a system of communications that determines the peculiar structure of these societies. There is no privacy. Information is locally shared.

5.2. The Nature of Signals Discovered by Benveniste

Dr. Benveniste discovered the existence of water memory, but this phenomenon could not yet be explained in 1988. Since his experimental results were attributed to error or fraud, he was constantly searching simpler and more objective methods to prove the reality of water memory. In 1990, he began to use the system of Langendorff ([6], p. 123). It is based on an allergic reaction of isolated hearts of guinea pigs or rats and modifies the coronary flow by about 20% ([10], pp. 337-343, 354). He used empirical methods, but had heard about the theory of Coherent Domains, where EM fields seemed to be involved. He tried thus to verify if *magnetic* fields of 50 Hz, easily produced by means of the European electric distribution system, might have some effect. About 100 experiments were performed in 1990 and 1991 with physicists of the Central Laboratory of Magnetism in France. It appeared that these fields could annihilate the biological efficiency for EHDs of histamine ([6], pp. 126-128).

This fact suggested that EM signals might be involved in water memory, but how could that be proven? Benveniste spoke in 1992 to a friend, who was an electronics hobbyist and thought that if molecules are able to produce EM waves, it might be possible to detect them by means of a *coil*. He constructed a kit, used for amplifying telephone sounds. It turned out that this method was sufficient to detect signals, created by an EHD that was contained in a tube, simply placed inside a coil. The wire delivered an *electric* signal that could be amplified ([6], pp. 128-132) by means of an electronic circuit that is valid for microphones, detecting audible sound waves. The detected signal varied thus at low frequencies (20 Hz - 20 kHz). When the amplified signal was applied to a second coil and when another tube, containing only pure water, was placed in this coil, it acquired the capacity to stimulate the same receptors than the original EHD. This fact did prove that EHDs produce "signals" that allow for information transfer. Beauvais called this a "molecular telephone" ([10], p. 336).

It was not possible anymore to attribute the published results to errors or sloppy work. Benveniste tried thus to restore his credibility in 1993 by means of experiments, performed with the system of Langendorff and controlled by physicists in Georges Charpak's laboratory in Paris. Charpak told Benveniste that if molecular communications were possible by means of ELF waves, that "would be the biggest discovery since Newton, if it were true". However, he was convinced that some fakers in Benveniste's laboratory did "arrange" the experimental results ([6], pp. 155-157). Nevertheless, it was decided that two types of molecules (acetylcholine and ovalbumin) should be used for preparing EHDs. They would be blindly treated by means of coding, but the results were catastrophic. For 11 of 18 series of tests performed in 1994, the reaction of the isolated hearts was inversed. Authentic EHDs had no biological efficiency anymore, while pure water produced positive results.

Benveniste tried to understand this fact. He knew that the tubes had not been exchanged, since he did transport them himself, but they had been placed side by side. The capacity of genuine EHDs to provoke biological reactions could thus have been transferred to pure liquid water, even without needing intermediate detection and amplification. To test this hypothesis, he did shield all samples by means of thick aluminum foil, but such a Faraday cage was not sufficient to suppress the unexpected information transfer. We conclude that it was not merely due to oscillating *electric* fields, but for Charpak, the idea of information transfer from tube to tube was even more abstruse than water memory. Actually, he wrote in July 1995 to Benveniste that he advances "the most baroque reasons to explain the failures" ([6], p. 166). This reaction was revolting for Benveniste. He was accustomed to attach primary importance to experimental evidence, but was now confronted with the conviction that "a phenomenon that cannot be explained with present-day knowledge cannot exist." ([6], p. 90).

Benveniste considered, on the contrary, that the new, objectively established facts demonstrated that *information transfer* is possible. It has to result from signals, which could be "intrinsic to molecular activity" ([6], p. 164). How this might be achieved remained mysterious, but Benveniste continued to improve his experimental methods ([13], pp. 336, 461, 519, 542, 584). His now very small laboratory acquired in 1995 a computer with a program for sound analysis. The electric signals, obtained by means of a coil could be digitalized before applying them to the second coil ([10], p. 461). Pure water became active, when it was exposed to this signal, even when it was digitally transmitted between different continents. He described these results in an article, submitted for publication to the *Journal of Immunology*. The final answer was that "the protocol and experiments are perfect, but the article cannot be published without precisions concerning the exact physical nature of the signal" ([6], p. 179).

It was only known that molecules can emit and absorb EM waves in the frequency domain of microwaves or infrared light. That molecules are able to emit signals of *very low frequencies* seemed to be impossible. Moreover, it was believed that biologically active molecules can only stimulate their receptors according to the model of *chemical reactions.* Since water memory transgressed this dogma, the well-known neurologist Changeux called it a "scientific heresy" ([6], p. 194). Why did referees and "experts" not verify if the reported facts are true? Why did the physicist Charpak not help to identify the nature of the signals? Benveniste recognized that "classical biologists" are not competent in electro-physics, since they use "structural models" ([6], pp. 180, 184). He mentioned also that he had only "adopted the hypothesis of an electromagnetic nature of the molecular signal" to organize empirical research ([6], p. 211). He was scientifically isolated and Charpak told him even: "your experiments challenge elementary laws of physics and common sense" ([6], p. 212). As far as we know, the nature of the relevant signals has never been identified. We summarize our explanation in **Figure 34**. The upper line corresponds to the description that we used so far. Electrically charged parts of active molecules do create an oscillating *electric* field.

We could compare the frequency f of the oscillating electric charge of active molecule to the *message* that has to be transmitted to potential receptors. This can be done in a direct way or by means of a bypass, as indicated in **Figure 20**. Trimmed chains are then the information carriers and could be compared to *messengers.* Anyway, the information has to be expressed or encoded in a particular way. For usual language, it could be encoded in oral or written form. For water memory, the *vehicle* can be an oscillating electric field, but this is not the only possibility. This results from the fact that oscillating electric charges are equivalent to electric currents. They create an oscillating *magnetic* field. [Initially, Oersted did use Volta's pile to show that motions of electric charges inside a wire do heat this conductor, but he noted that the current caused a rotation of a nearby compass needle. This phenomenon was totally unexpected, since it was only known that magnetic materials are able to do that. Oersted described in 1820 the properties of the new type of "magnetic forces" in terms of actions at a distance, as Newton did for gravitational forces and Coulomb for electric forces.]

The concepts of electric and magnetic fields were introduced by *Faraday.* [He discovered in 1831 that imagined "lines of magnetic forces" passing through a closed loop of a conductive wire induce there an electric current, but only when the flux is varying. It can increase, decrease or oscillate. In 1845, he used the

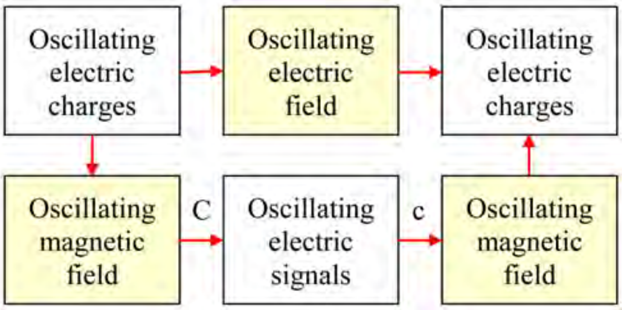


Figure 34. Communications by means of oscillating electric and magnetic fields inside EHDs, with the aid of two coils (C and c).

more general concept of electric and magnetic fields, defined for any point in space and time by means of a fictitious experiment. These fields can then vary in space and time.] The second line of **Figure 34** shows that oscillating *electric* fields can be replaced by oscillating *magnetic* fields. For Benveniste and Montagnier these fields were hidden in "black boxes". They become now mentally transparent and even brilliant. Charpak could have explained what was in these black boxes, but did not even care, since they were irrelevant for him.

The second line of **Figure 34** indicates also that when a tube with an EHD of a particular type of active molecules is placed inside a coil C, the oscillating *magnetic* field exerts a force on charged particles in the wire. This force produces a current when the circuit is closed or a tension between the extremities of an open circuit. [This fact results from Faraday's discovery.] The resulting *electric* signal can be amplified and applied to a second coil c. It produces a current in the wire of this coil and thus again a *magnetic* field. It oscillates at the same frequency and exerts a force on electric charges. Benveniste compared the relevant signals to radio waves. This analogy was helpful, but they were merely electric and magnetic fields, oscillating at low frequencies. They require the presence of electric charges, while EM waves consist of electric and magnetic fields that are constantly transformed into one another and can even propagate in vacuum. They are characterized by the relation $f = c/\lambda$, where the velocity $c \approx 3 \times 10^8$ m/s.

5.3. Stimulation by Low-Frequency Magnetic Fields

After realizing that water memory is transferable to pure water, Benveniste developed "digital biology" with the engineer *Didier Guillonnet,* who joined the team in 1996 ([13], p. 596). Their common patent [99] indicates that information transfer is even possible when pure water is simply flowing through the second coil (c). This means that the formation of new trimmed chains of WPs is a rapid process. It would even be sufficient to create one trimmed chain of WPs, since it will generate other equally trimmed chains of WPs. This fact is also important for beat phenomena, since equally trimmed chains of WPs are then reproduced inside the initially pure water. Benveniste had noted that shaking of the receiving tubes was not required, but Benveniste and Guillonnet found that "it is possible to improve the performance" of this system by adding a signal that corresponds to *white noise.* It covered the whole low frequency domain (1 Hz to 20 kHz). They mentioned that they were surprised when they discovered this fact. Konovalov and Ryzhkina noted also that the formation of supramolecular structures is favored by low-frequency EM fields [87].

Montagnier and his collaborators [14] placed the test tubes, which contain "informed" and "naive" water in different coils (C and c), but inside a greater coil (C'). It was subjected to an electric current, *oscillating at* 7 Hz*.* This frequency was chosen since it is close to "Schumann frequencies". They correspond to genuine *EM waves*, generated by thunderstorms in the terrestrial atmosphere and reflected by the surface of the Earth and the ionosphere. Since these waves are propagating in a spherical layer, they correspond to standing waves and their wavelength λ is of the order of the circumference of the Earth. It follows that their frequency $f = c/\lambda \approx 300$ (thousand km/s)/40 (thousand km) = 7.5 Hz. The measured fundamental frequency for Schuman resonances is close to 7.8 Hz. However, the coil C' creates only an oscillating *magnetic* field and not an EM field, similar to radio waves of extra low frequencies.

The term of "electromagnetic fields" was probably used by Benveniste and Montagnier in the general sense of being related to electric and magnetic phenomena. To realize the difference, we note that standing waves on trimmed chains of WPs could be compared to standing waves for oscillating electrons in an antenna that radiates EM waves. Their wavelength λ would then be determined by the length of these chains ($L = \lambda/2 \approx 300$ nm), but their frequency would be extremely high ($f = c/\lambda \approx 5 \times 10^{14}$ Hz). The inverse process corresponds to an EM wave that excites standing waves for oscillating electrons in a receiving antenna. For 7 Hz, the wavelength $\lambda = c/f \approx 43{,}000$ km.

Nevertheless, a coil C' and low frequency currents were beneficial. To explain this fact, we recall that water molecules are bound to one another in bulk liquid water by hydrogen bonds (**Figure 4(a)**), but are constantly exchanged at the time scale of 10^{-12} s. The applied magnetic field is oscillating at a much lower frequency (7 Hz or 1 Hz - 20 kHz) and does reorient the effective dipoles of water molecules at this frequency. They will thus be liberated from their usual bonds. This facilitates the formation of water pearls and trimmed chains of WPs.

We recall also that "wild transfers" were not suppressed when the test tubes were shielded by means of aluminum foil. It did only suppress electric fields. That was not sufficient, since magnetic fields do also allow for information transfer. However, it is possible to eliminate magnetic fields by means of mu-metal. Benveniste had already discovered that it abolishes crosstalk and beat phenomena ([10], p. 565) and Montagnier [14] did interpose a sheet of mu-metal between the primary coil C and the secondary coil c. Empirical discoveries do not require that their cause is already known.

5.4. Resonances for Pairs and Chains of Water Pearls

We come now back to the important and still unsolved problem of the *asymmetry* of the energy-loss functions for Mahata's resonances (**Figure 33**). The model of single WPs has to be corrected, since WPs would be subjected to Brownian motion. They attract one another and will thus constitute *pairs* of WPs. Oscillatory rotations are then possible for both WPs, even at a frequency $\Omega \approx 25$ MHz), because of surrounding water molecules. However, mutual attraction of neighboring poles would yield an additional restoring force, which could have noticeable effects.

To acquire more physical insight, we consider two identical pendulums (**Figure 35(a)**). They would oscillate at the same frequency Ω, but can be coupled by means of a weak spring. [A normal string is sufficient, when it is somewhat longer than the separation of the masses at rest and when a small mass is attached to

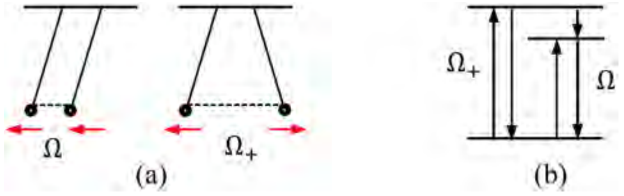


Figure 35. (a) Low and high frequency modes for free oscillations of two coupled pendulums; (b) Resonances correspond to quantum mechanical transitions, allowing for excitation at the frequency Ω_+ and emission at the frequency Ω.

the middle of the string.] Free motions of the coupled pendulums depend on the chosen initial conditions and can be quite complicated, but there are two special cases, shown in **Figure 35(a)**. It is intuitively clear that both masses will oscillate at the frequency Ω for strictly parallel motions, since the length of the intermediate spring is not modified. It is higher for antiparallel motions ($\Omega_+ > \Omega$). These coordinated motions are called *normal modes of oscillation,* since the system behaves like a single oscillator. It is possible to produce forced motions of both pendulums, by applying an oscillating force to one of them.

Figure 35(b) recalls that QM is a generalization of classical mechanics. The lowest possible energies for any single harmonic oscillator are then separated by intervals that are proportional to the resonance frequency. This can be Ω or Ω_+. Classically described forced oscillations are then replaced by *transitions*. It should thus be possible to absorb EM energy at the frequency Ω_+ and to restitute a part of it at the lower frequency Ω. This happens for fluorescence and Mahata's energy-loss functions $\beta_2(\omega)$ displayed always at the side of lower frequencies (**Figure 33**).

Figure 36 shows that we get also two normal modes for small oscillatory rotations of the coupled WPs. As for **Figure 35(a)**, displacements of neighboring poles can be parallel or antiparallel. They are perpendicular to the symmetry axis and represented by red arrows. To account for possible resonances, we have to consider two coupled equations of motion:

$$\ddot{u}_1+\Omega^2u_1+\omega_o^2\left(u_1-u_2\right)+\gamma_1\dot{u}_1=C_\omega\exp\left(-i\omega t\right) \tag{19}$$

$$\ddot{u}_2+\Omega^2u_2-\omega_o^2\left(u_1-u_2\right)+\gamma_2\dot{u}_2=-C_\omega\exp\left(-i\omega t\right) \tag{20}$$

They apply to the left part of **Figure 36** and Ω ≈ 25 MHz is the resonance frequency for independent oscillations, while $\omega_o \approx 1$ kHz accounts for interactions. They correspond to opposite forces. For the right part of **Figure 36**, it is sufficient to change the sign of u_2. Both water pearls are subjected to viscous friction, but it is not identical, since the resulting motions of the surrounding liquid water can hinder or facilitate one another. To assume that $u_1 > u_2$ implies that $\gamma_1 < \gamma_2$. When we set $\gamma_1 = \gamma - \eta$ and $\gamma_2 = \gamma + \eta$, the sum and the difference of (19) and (20) yield two equations for $u_\pm = u_1 \pm u_2$:

$$\ddot{u}_+ +\Omega^2u_+ +\gamma\dot{u}_+ -\eta\dot{u}_- =0 \tag{21}$$

$$\ddot{u}_- +\Omega_+^2u_- +\gamma\dot{u}_- -\eta\dot{u}_+ =2C_\omega\exp\left(-i\omega t\right) \tag{22}$$

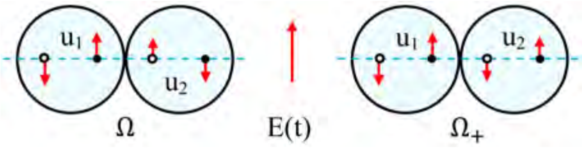


Figure 36. Two modes of forced motions for pairs of WPs.

The pair of WPs behaves thus as if there were two oscillators that resonate at the frequencies Ω or Ω_+, where $\Omega_+^2 = \Omega^2 + 2\omega_o^2$. However, only the mode u_- is directly excited by the oscillating electric field of angular frequency *ω*. The mode u_+ is then excited by *entrainment*, but forced oscillations require that

$$u_\pm = A_\pm 2C_\omega \exp(-i\omega t)$$

Because of (21) and (22), the amplitudes $A_\pm$ are determined by the equations:

$$a_+ A_+ + i\omega\eta A_- = 0 \quad \text{and} \quad a_- A_- + i\omega\eta A_+ = 1$$

where $a_+ = \Omega^2 - \omega^2 - i\omega\gamma$ and $a_- = \Omega_+^2 - \omega^2 - i\omega\gamma$. Thus,

$$A_+ = -\frac{i\omega\eta}{a_+} A_- \quad \text{and} \quad \left[a_- + \frac{(\omega\eta)^2}{a_+} \right] A_- = 1$$

Forced oscillatory rotations of pairs of WPs involve both modes of oscillation. The polarizability is thus

$$\beta(\omega) = \Omega_+^2 (A_- + A_+)$$

It is normalized to get *β*(0) = 1. When Ω and ω_o correspond to 25 MHz and 1 kHz, there are only two adjustable parameters: *γ* and *η*. They specify the average energy loss and energy transfer by viscous friction. The resulting values of $a_\pm$ and $A_\pm$ allow us to calculate the spectral distribution of the real and imaginary parts of *β*(*ω*). When $\gamma = 20$ and $\eta = 0.01$ (MHz), the calculated spectral distributions of $\beta_1(\omega)$ and $\beta_2(\omega)$ are shown in **Figure 37**. The energy-loss function presents a dip, like the experimentally observed one (**Figure 33**). The concept of water pearls accounts thus also for the asymmetry for pairs of WPs.

However, Mahata found similar curves for EHDs of various types of active molecules. This would not be compatible with the excitation of a standing wave, if all positive poles were displaced along the direction of the applied electric field E. However, the lowest possible excitation energy would only require that the rotation of *one* water pearl is inversed. **Figure 38** shows an example. Even the excitation of standing waves on trimmed chains of WPs at high frequency and the asymmetric resonances can thus be explained by the proposed theory.

5.5. The Stabilizing Effect of Lactose

Homeopathic preparations are often presented in the form of pills of lactose. Opponents of this medical practice claim that lactose is merely used because of its sweet taste. Initially, Hahnemann used lactose for grinding hard substances to reduce their concentration before dissolving them in pure liquid water. Mahata

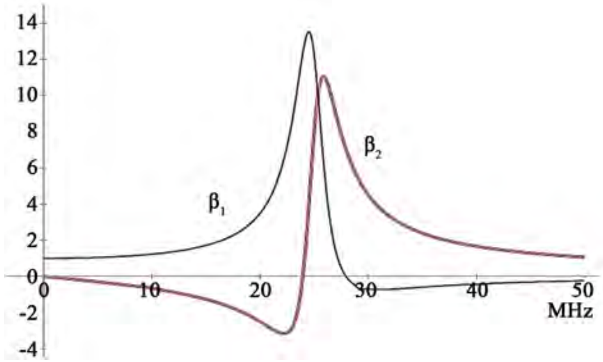


Figure 37. Predicted spectral distribution of the realpart (β_1) and the imaginary part (β_2) of the polarizability of pairs of coupled water pearls in pure water.

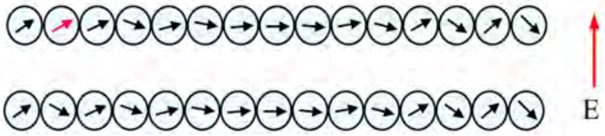


Figure 38. Resonance for trimmed chains of WPs.

knew that lactose is also used to insure better preservation of the biological efficiency of EHDs. He wanted thus to find out if the association of EHDs with lactose modifies the measurable resonance curves. When he dissolved lactose powder in pure water, he found a resonance at 50 MHz and the energy-loss function was symmetric [100]. We have thus also to explain these facts. From the point of view of condensed matter physics, lactose is an insulator. It attracts thus charged particles by *image forces.* Since lactose has many pores, nearly all single WP would be attached to lactose, but the applied electric field would simultaneous set the WP and its image in oscillation. They are subjected to identical restoring forces. The resonance frequency will thus be doubled, but attachment of single WPs would keep the resonance symmetric, as for the red curve in **Figure 33**.

When Mahata added EHDs of Cu-Met-30 to the dissolved lactose powder, the resonance frequency was reduced to about 45 MHz, but the peak of the energy-loss function remained symmetric. Because of image forces, the trimmed chains of WPs would usually be parallel to the surface of the insulator. This will also lead to a high resonance frequency, but its value would be somewhat lower than 50 Hz. The essential result is, of course, that the protective role of lactose is due to adhesion by image forces.

5.6. A New Type of Molecular Interactions and Medical Applications

The empirical discovery of water memory and its elucidation modify the traditional paradigm that molecular interactions are only possible according to the "key and slot" model of chemical reactions. Even normal interactions between biologically active molecules and their specific receptors are due to oscillating electric fields and resonance effects (upper lines of **Figure 20** and **Figure 34**). Dr. Hahnemann wanted only to select medications in a rational way and to increase safety of medical practice. Dr. Benveniste did prove that this is due to the phenomenon of water memory. He asked INSERM several times that "interdisciplinary teams" could come to help him to interpret what he observed. Scientific institutions should have provided intellectual and materiel assistance, instead of "isolating" him. Benveniste felt that these authorities did even "strangle" him ([10], p. 89). The case of water memory has to be remembered as a monument of what should not happen, especially in science.

Moreover, it is not unusual that new phenomena are discovered without understanding their cause. We mentioned many examples that illustrate this fact. *Christian Huyghens* discovered in 1665 that pendulum clocks tend to be synchronized, but this phenomenon has only recently been explained [101]. Collective oscillations are also very important in plasma physics and provide for instance the key for understanding the occasional appearance and peculiar properties of ball lightning [102]. Oscillations of the charges of biologically active molecules and standing waves on trimmed chains of water pearls create electric and magnetic fields that tend also to be synchronized. This favors the autocatalytic generation of new trimmed chains and does more effectively stimulate the relevant receptors.

Since oscillating electric charges do also produce oscillating *magnetic* fields (**Figure 34**), we could try to pursue investigations by means of *modern magnetometers*. Squids are used for magneto-encephalography, but a new, less expensive and very efficient technology is now available in the form *spintronics.* This word is an abbreviation of spin-electronics. Instead of exerting forces on the electric charge of electrons, it is indeed possible to exert also forces on their spin. This technology has medical applications [103]. New magnetic detectors of this type have a very high sensitivity in a large frequency domain (in the picotesla region for frequencies below 100 Hz). Their small active areas allow for unprecedented spatial resolution (down to tens of nanometers). They seem thus to be ideal tools for trying to study interactions between active molecules and their receptors. They could open new avenues for future research in biology, medicine and pharmacology, if it were possible to detect low resonance frequencies of biologically active molecules and their specific detectors.

This has already been achieved by means of EHDs and there are observations that indicate the usefulness of such measurements. The collaboration of Mahata with the medical doctor Chattopadhyay, an Indian specialist of homeopathy, led indeed to are markable discovery [104] [105]. They found that EHDs of the blood serum of patients did also produce resonances at about 25 MHz. They were similar to those of EHDs of biologically active molecules. This proves that they have charged parts that can also oscillate and create trimmed chains of WPs. Moreover, Chattopadhyay found that for patients, having ailments that can be treated by means of homeopathic medicines, preparations that led to a closer match of resonance curves were more efficient. It was concluded that it should be possible to "make a *selection of medicines* based on scientifically measurable parameters." The reported results require independent confirmations, but it would not be prudent to discard this possibility without further objective studies.

The basic claim of Montagnier's patent US2010323391 [47] was that hidden infections in humans and animals can be detected by a non-invasive method. It requires only to collect blood samples and to prepare EHDs of them. This allowed for spectral analysis of signals that were detected by means of coils. Another patent added that "any human or animal fluid, e.g. blood, urine, various secretions" can be used for this purpose. Even the unexplained resurgence of some sicknesses might result from water memory, since Montagnier [14] [15] found that DNA molecules do also produce similar signals.

Another potentially important question concerns Hahnemann's empirical rule that "like cures like". Is it possible to prove the existence of a link between receptors and the sickness they can cure? Present-day knowledge and already acquired experience in the domain of receptors and neurology could be used, for instance, to verify if some sicknesses are related to particular receptors, by exiting them. Research is motivated by curiosity, whatever may be found.

6. Conclusions

The objective of this study was to find out if water memory is possible or not. We treated this problem in terms of condensed matter physics, but it illustrates a much more general and fundamental difficulty: the recurrent *conflict between facts and ideology.* In science, it is not unusual that empirical research uncovers phenomena without understanding the underlying mechanism. It does even happen that the framework of existing theories cannot explain them. Past experience and commonsense tell us that previous assumptions may have to be corrected in such a situation. Kuhn has shown [45] that the discovery of an anomaly with respect to "normal theory" does often lead to initial rejection. This was also true for water memory. It is thus necessary to be aware of the fact that scientific progress can be drastically slowed down by so-called "skeptics", who are not skeptic enough in regard to their own preconceptions and beliefs.

It is true that "extraordinary claims require extraordinary evidence," but this slogan does only displace the basic problem. What is valid evidence? In science, it can merely be recognized by referring to *reality.* This is well-known, but unfortunately, there is a strong tendency to rely only previously acquired ideas and theories, although they could have been based on hypotheses that were only valid in a limited domain. They may have to be replaced by more general ones. The case of water memory illustrates such a need in a rather exemplary way.

Physics is also confronted today with a similar problem. It results indeed from observations that there are only certain types of elementary particles and that our Universe contains an enormous amount of Dark Matter. The accelerated expansion of space is caused by Dark Energy, but we are unable to explain all these facts. It is thus necessary to ask if present-day theories do not contain an assumption that was simple and useful, but merely an approximation. In this regard, we learned even from the development of the theory of relativity and quantum mechanics that Nature can impose restrictions on our measurements. It appeared that they are related to the existence of two universal constants (c and h). Nevertheless, we continue to believe that space and time are continuous. This is equivalent to postulating that there is no finite limit for the smallest measurable distance. How do we know? We can only say that until now, we did not yet meet such a limit. However, we can try to find out what would happen if there did exist a universally constant quantum of length (a) and thus also a universally constant quantum of time (ca).

The value of a is surely very small, but we cannot assert that $a = 0$. If this value were finite, all physical laws for particles and fields that involve variations in space and time would not be expressed anymore by differential equations, but by finite-difference equations. [The differential wave Equation (7) was also an approximation of the more general finite-difference Equation (6), but for another reason.] When we did that for any type of particle and force fields, it turned out that the generalization would not lead to logical inconsistencies when $a \neq 0$. However, some ideas have to be changed. The highest possible energy, which has to be attributed to the whole Universe, would be finite. The behavior of fields at the smallest possible scale in space and time would be described by hitherto unknown quantum numbers. They account for all possible types of elementary particles, in agreement with already known facts [106]. This applies also to Dark Matter. It is known to be present in our Universe, but is composed of particles that have not yet been identified by means of accelerators. In cosmology, it is also possible to account for the accelerated expansion of space and Dark Energy [107]. This enlargement of the foundations of physics widens the domain of possibilities, but the concept of a "space-time continuum" is so deeply rooted in our minds and our culture, that it will take time to be able to modify this idea.

This is also true for the conviction that biologically active molecules can only interact with their specific receptors according to the model of chemical reactions. The possible existence of water memory seemed to be absurd, but the real problem was merely that it could not yet be explained. Of course, ferroelectric crystallites of water molecules and trimmed chains of water pearls are merely concepts. These entities are not directly observable, but they allow us to make verifiable predictions. An increasing part of science follows this pattern. Theory and experiment are complementary. On one hand, we have to imagine what might be possible, to draw logical consequences from the proposed hypotheses. On the other hand, we can establish what it real or not. Sometimes the observations precede their explanation and sometimes, they can be used to test the validity of hypothesis and theories.

In regard to water memory, we found that the concept of trimmed chains of WPs accounts for the quasi-periodic variations of the biological efficiency of EHDs (Sections 3.4 and 3.5), the measurable frequency spectrum (Section 4.1) and the peculiar angles for junctions of large-scale structures (Section 4.2). It is at least probable that very long chains of WPs explain the stability of liquid water bridges (Section 4.3). The von Grotthuss mechanism can be understood in terms of intramolecular exchange effects for delocalized protons (Section 2.5). They are also relevant for 2D polymerization of water molecules (Section 2.6). The perplexing effects of cross talk can be attributed to beat phenomena (Section 3.7). Physicochemical and other types of measurements make sense (Sections 3.8 and 4.5). It is also possible to understand the physical nature of detected signals (Sections 5.2 and 5.3). Even Mahata's unexpected high frequency resonances can be explained (Sections 4.6, 5.4 and 5.5).

Of course, there are still open questions, inviting to pursue research (Section 5.6). We wonder for instance if low-frequency resonances can be detected for single chains of WPs. Preliminary results for dielectric responses of EHDs indicated that a resonance could be excited at about 3 kHz [108]. Miranda did perhaps not detect them [74], since that would have required greater initial concentrations of LiCl or various types of biologically active molecules. The now highly developed methods of X-ray and neutron scattering could also be applied to EHDs of initially very great concentrations of ions or active molecules. We mention that Mahata obtained images of elongated entities that constituted a super lattice. These structures were observed by scanning probe microscopy [100]. Can variations of the distance between oxygen ions near the surface of water bridges subjected to different high tensions be measured diffraction of X-rays? Can oscillating magnetic fields be detected at molecular levels by means of modern magnetometers? Could water memory have contributed to the emergence of life on Earth?

Water is a very familiar substance, but still a fascinating domain of research. Moreover, it concerns not only experimental and theoretical results, but also truth and justice.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Davanas, E., Beauvais, F., Amara, J., Oberbaum, M., Robinzon, B., Miadonnai, A., Benveniste, J., *et al.* (1988) *Nature*, **333**, 816-818. https://doi.org/10.1038/333816a0

[2] Maddox, J., Randi, J. and Stewart, W.W. (1988) *Nature*, **334**, 287-291. http://www.badscience.net/wp-content/uploads/benveniste02.pdf

[3] Maddox, J. (1988) *Nature*, **333**, 787. https://www.nature.com/articles/333787a0 https://doi.org/10.1038/333787a0

[4] Benveniste, J., Henson, P.M. and Cochrane, C.G. (1972) *Journal of Experimental Medicine*, **136**, 1356-1377. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2139324/ https://doi.org/10.1084/jem.136.6.1356

[5] Benveniste, J., Egido, J. and Gutierrez-Millet, V (1976) *Clinical & Experimental*

Immunology, **26**, 449-456.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1540955/

[6] Benveniste, J. (2005) Ma vérité sur la « mémoire de l'eau. Book, Albin Michel, 223 p.

[7] Baldwin, M. (2015) The History of Medicine in Dialogue with the Present, Remedia.
https://remedianetwork.net/2015/04/24/high-dilution-homeopathy-and-the-purpose-of-the-scientific-journal/

[8] Kaufmann, A. (1994) *Réseaux*, **2**, 183-204.
http://www.persee.fr/doc/reso_0969-9864_1994_num_2_2_3277

[9] Schiff, M. (1994) Un cas de censure dans la science: L'Affaire de la mémoire de l'eau, AlbinMichel. (1995); The Memory of Water, Thorsons, Paperback, 166 p.

[10] Beauvais, F. (2007) L'âme des molécules, une histoire de la "mémoire de l'eau", Book, Mille-Mondes; (2016) Ghosts of Molecules—The Case of the "Memory of Water". Ed. Lulu.com.

[11] Thomas, Y. (2007) *Homeopathy*, **96**, 151-157.
https://www.ncbi.nlm.nih.gov/pubmed/17678810
https://doi.org/10.1016/j.homp.2007.03.006

[12] CristianManil (2014) On a retrouvé la Mémoire de l'eau, Luc Montagnier, TV, 51.42 min. https://www.youtube.com/watch?v=_2xInJFD23k

[13] Montagnier, L. (2016) Water Memory.
https://www.youtube.com/watch?v=R8VyUsVOic0

[14] Montagnier, L., Aïssa, J., Ferris, S., Montagnier, J.-L. and Lavallée, C. (2009) *Interdisciplinary Sciences*, **1**, 81-90. https://doi.org/10.1007/s12539-009-0036-7

[15] Montagnier, L., *et al.* (2010) DNA Waves and Water.
https://arxiv.org/pdf/1012.5166.pdf

[16] Montagnier, L., *et al.* (2014) Digital Transmission of Bacterial DNA in Living Cells, UNESCO, Biology in Light of Physical Theories: New Frontiers in Medicine.
https://www.youtube.com/watch?v=mYcNXjdPXGU

[17] Chaplin, M.F. (2007) *Homeopathy*, **96**, 143-150.
http://www.doktor-quak.de/wp-content/uploads/2014/01/The-Memory-of-Water-an-overview.pdf

[18] Chaplin, M. (2016) Water Structure and Science.
http://www1.lsbu.ac.uk/water/water_molecule.html

[19] Chaplin, M. (2016) Water Models.
http://www1.lsbu.ac.uk/water/water_models.html

[20] https://en.wikipedia.org/wiki/Kinetic_diameter

[21] Rouard, P. and Meessen, A. (1977) *Progress in Optics*, **VX**, 77-137.
http://www.meessen.net/AMeessen/optical_properties_of_thin_metal_films.pdf
https://doi.org/10.1016/S0079-6638(08)70477-3

[22] Granqvistt, C.G. and Buhrman, R.A. (1976) *Journal of Applied Physics*, **47**, 2200-2219.
https://buhrman.research.engineering.cornell.edu/pubs/JApplPhys_47_2200.pdf
https://doi.org/10.1063/1.322870

[23] Latimer, W.M. and Rodebush, W.H. (1920) *Journal of the American Chemical Society*, **42**, 1419-1433.
http://www.chemteam.info/Chem-History/Hydrogen-Bond-1920/Hydrogen-Bond-1920.html
https://doi.org/10.1021/ja01452a015

[24] Stryer, L. (1981) Biochemstry. Stanford University Freeman and Co., Text Book, 2nd Edition, 125.

[25] Chaplin, M. (2016) Ice Phases. http://www1.lsbu.ac.uk/water/ice_phases.html

[26] Mallamace, F. (2009) *PNAS*, **116**, 1597-1598.
http://www.pnas.org/content/106/36/15097
https://doi.org/10.1073/pnas.0908198106

[27] Mukhopadhyway, A., *et al.* (2015) *Chemical Physics Letters*, **623**, 13-26.
http://www.cchem.berkeley.edu/rjsgrp/publications/papers/2015/401_Mukhopadhyay.pdf
https://doi.org/10.1016/j.cplett.2015.04.016

[28] Köffinger, J., *et al.* (2011) *Physical Chemistry Chemical Physics*, **13**, 15403-15417.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3470881/

[29] Stillinger, F.H., *et al.* (1978) *The Journal of Chemical Physics*, **69**, 1473-1484.
https://pdfs.semanticscholar.org/6c11/c7b39c3e2b02eb65c27e0a5600875c9128b6.pdf
https://doi.org/10.1063/1.436773

[30] Von Grotthus, T. (1806) *Annales de Chimie* (*Paris*), **58**, 54-73.

[31] Lippincot, E.R. and Stromberg, R. (1969) *Science*, **164**, 1482-1487.

[32] Stromberg, J. (2013) The Curious Case of Polywater.
http://www.slate.com/articles/health_and_science/science/2013/11/polywater_history_and_science_mistakes_the_u_s_and_ussr_raced_to_create.html

[33] Roy, R., *et al.* (2009) *Materials Research Innovations*, **9**, 98-103.
https://doi.org/10.1080/14328917.2005.11784911

[34] Pollack, G. (2009) Water, Energy, and Life: Fresh Views from the Water's Edge.
https://www.youtube.com/watch?v=XVBEwn6iWOo

[35] Kittel, C. (1946) *Physical Review*, **70**, 965-971.
https://pdfs.semanticscholar.org/8639/78264c8d8b616d37b96a9147b65595b61fdd.pdf
https://doi.org/10.1103/PhysRev.70.965

[36] Blakemore, R.P. (1975) *Science*, **190**, 377-379.
https://doi.org/10.1126/science.170679

[37] Faivre, D. and Schüller, D. (2008) *Chemical Reviews*, **108**, 4875-4898.
http://www.cup.lmu.de/ac/kluefers/homepage/L/biominerals/magnetite_2008.pdf
https://doi.org/10.1021/cr078258w

[38] Eliaseev, E.A., *et al.* (2016) *Journal of Applied Physics*, **119**, 204104.
https://arxiv.org/abs/1602.06285
https://doi.org/10.1063/1.4952707

[39] Smith, A.M., *et al.* (2016) The Electrostatic Screening Length in Concentrated Electrolytes Increases with Concentration.
https://arxiv.org/ftp/arxiv/papers/1607/1607.03926.pdf

[40] Lo, S.Y. (1998) Physical, Chemical and Biological Properties of Stable Water Clusters. *Proceedings of the First International Symposium*.
http://dandylaboratories.com/wp-content/uploads/2011/09/SURVEY-OF-I-CLUSTERS.pdf

[41] HNMRC (2015) Information Paper. Evidence of the Effectiveness of Homeopathy for Treating Health Conditions, Australian Government.
https://www.homeowatch.org/research/nhmrc_2015.pdf

[42] Homeopathy Research Institute (2012).
http://www.nhmrchomeopathy.com/the-first-homeopathy-review.html

[43] Shurkin, J. (2014) News Feature: Animals that Self-Medicate.
http://www.pnas.org/content/111/49/1733
https://doi.org/10.1073/pnas.1419966111

[44] Hahnemann's "Organon of Medicine" (1833, 1842) Transl. by R.E. Dudgeon and W. Boericke. http://www.homeopathyhome.com/reference/organon/organon.html

[45] Kuhn, T.S. (1962) The Structure of Scientific Revolutions. University of Chicago Press. (1st Edition, 172 p; 2012, 4th Edition, 264 p).

[46] Elia, V and Nicoli, M. (1999) *Annals of the New York Academy of Sciences*, **879**, 241-248. https://doi.org/10.1111/j.1749-6632.1999.tb10426.x

[47] Elia, V., *et al.* (2004) *Journal of Thermal Analysis and Calorimetry*, **78**, 331-342.
https://doi.org/10.1023/B:JTAN.0000042179.07858.c8

[48] Elia, V., *et al.* (2007) *Homeopathy*, **96**, 163-169.
http://omeopatia.org/download/seminario-rey/ELIA-Dell-acqua-la-rimembranza-inglese.pdf
https://doi.org/10.1016/j.homp.2007.05.007

[49] Elia, V., *et al.* (2008) *Journal of Thermal Analysis and Calorimetry*, **93**, 1003-1011.
https://doi.org/10.1007/s10973-007-8843-8

[50] Elia, V., *et al.* (2014) Conference on the Physics, Chemistry and Biology of Water.
https://www.youtube.com/watch?v=EPVHl3xR7p0

[51] Montagnier, L. (2011) DNA "Teleportation".
http://www.rexresearch.com/montagnier/montagnier.htm

[52] Alm, J.F. and Walker, J.S. (2002) *SIAM Review*, **44**, 457-476.
https://epubs.siam.org/doi/pdf/10.1137/S00361445003822
https://doi.org/10.1137/S00361445003822

[53] Lo, S.Y. (1996) *Modern Physics Letters B*, **10**, 921-930.
https://www.cs.cmu.edu/~dst/ATG/lo-iestru.html
https://doi.org/10.1142/S0217984996001048

[54] Lo, S.Y., *et al.* (2009) *Physics Letters A*, **373**, 3872-3876.
https://www.researchgate.net/publication/222823593_Evidence
https://doi.org/10.1016/j.physleta.2009.08.061

[55] Gann, D. and Lo, S.Y. (2009) Double Helix Water. D and W Publ., 316 p.

[56] Ho, W.M. (2014) *Water*, **6**, 1-12. http://www.waterjournal.org/volume-6/ho

[57] Armstrong, W.G. (1893) *The Electrical Engineer*, **10**, 154-152.
http://ttm-old.tugraz.at/jw/down/Armstrong_1893.pdf

[58] Fuchs, E.C., *et al.* (2007) *Journal of Physics D: Applied Physics*, **40**, 6112-6114.
https://doi.org/10.1088/0022-3727/40/19/052

[59] Fuchs, E.C. (2016) Protons and the Floating Water Bridge (Interview at the Conference on the Physics Chemistry and Biology of Water).
https://www.youtube.com/watch?v=v8ES0Y0564M

[60] Fuchs, E.C. (2014) Protons and the Electrical Conduction in a Floating Water Bridge. Conference on the Physics, Chemistry and Biology of Water.
https://www.youtube.com/watch?v=ZB1bCt9JmeE

[61] Fuchs, E.C., *et al.* (2016) *Journal of Physics D*, **49**, 1-7.
http://iopscience.iop.org/article/10.1088/0022-3727/49/12/125502/pdf
https://doi.org/10.1088/0022-3727/49/12/125502

[62] Ponterio, R.C., *et al.* (2011) *Journal of Physics D*, **43**, 175405-175412.
http://iopscience.iop.org/article/10.1088/0022-3727/43/17/175405

https://doi.org/10.1088/0022-3727/43/17/175405

[63] Skinner, L.B., *et al.* (2012) *PNAS*, **109**, 16463-16468.
http://www.pnas.org/content/109/41/16463.full.pdf
https://doi.org/10.1073/pnas.1210732109

[64] Fuchs, E.C., *et al.* (2009) *Journal of Physics D*, **42**, 105502 (10 p).
https://hal.archives-ouvertes.fr/hal-00569550/document

[65] Namin, R.M., *et al.* (2013) *Physical Review E*, **88**, 033019 (6 p).
https://arxiv.org/abs/1210.2913

[66] Fuchs, E. (2009) The Inner Structure of a Floating Water Bridge, SETI Talks.
https://www.youtube.com/watch?v=N1At3Gcd-No

[67] Del Guidice, E., Preparata, G. and Vitiello, G. (1988) *Physical Review Letters*, **61**, 1085-1088.
https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.61.1085
https://doi.org/10.1103/PhysRevLett.61.1085

[68] Preparata, G. (1995) QED Coherence in Matter. World Scientific, 252 p.
https://doi.org/10.1142/2738

[69] Del Guidice, E. (2014) *Water*, **6**, 61-71.
http://www.waterjournal.org/uploads/vol6/delgiudice/WATER.2014.3.DelGiudice.pdf

[70] Del Guidice, E., *et al.* (2011) *Journal of Physics*: *Conference Series*, **329**, 012001.
http://iopscience.iop.org/article/10.1088/1742-6596/329/1/012001/pdf
https://doi.org/10.1088/1742-6596/329/1/012001

[71] Bellavite, P., *et al.* (2013) *Homeopathy*, **103**, 4-21.
https://www.researchgate.net/publication/259805970
https://doi.org/10.1016/j.homp.2013.08.003

[72] Montagnier, L., *et al.* (2017) *Water*, **9**, 339-357.
https://arxiv.org/pdf/1804.02436.pdf
https://doi.org/10.3390/w9050339

[73] Yinon, T.A. and Liu, Z.Q. (2015) *Water*, **7**, 33-95.

[74] Miranda, A.R., *et al.* (2011) *Materials Research Innovations*, **15**, 302-308.
http://www.fap.if.usp.br/~vannucci/vannucci_H2Oarticle.pdf
https://doi.org/10.1179/143307511X13109310554445

[75] Maity, T., Ghosh, D. and Mahata, C.R. (2010) *Homeopathy*, **99**, 99-103.
https://www.ncbi.nlm.nih.gov/pubmed/20471612
https://doi.org/10.1016/j.homp.2009.10.004

[76] Mahata, C.R. (2013) *Homeopathy*, **102**, 262-3267.
http://daneshyari.com/article/preview/2630007.pdf
https://www.sciencedirect.com/science/article/pii/S1475491613000635?via%3Dihub
https://doi.org/10.1016/j.homp.2013.07.003

[77] Mahata, C.R. (1997) *The Homeopathic Heritage*, **22**, 245-252.

[78] Barnard, G.P. (1965) *Journal of the American Institute of Homeopathy*, **58**, 205-212.

[79] Barnard, G.P. and Stephenson, J.H. (1967) *Journal of the American Institute of Homeopathy*, **60**, 277-286.

[80] Gay, A. and Boiron, J. (1953) Démonstration physique de l'existence réelle du remède homéopathique, Ed. Lab. P.H.R., Lyon, France (Communication Internal Congress in Brussels).

[81] Maity, T., Ghosh, D. and Mahata, C.R. (2007) *J. Inst. Electr. Enineers* (*India*), **88**, 27-31.

[82] Rey, L. (2003) *Phyica A*, **323**, 67-74. http://www.homresearch.org/docs/Physica%20A.PDF

[83] Rey, L. (2007) *Homeopathy*, **96**, 170-174. https://doi.org/10.1016/j.homp.2007.05.004

[84] van Wijk, R., *et al.* (2006) *Journal of Alternative and Complementary Medicine*, **12**, 437-443. http://www.doktor-quak.de/wp-content/uploads/2014/01/Thermoluminescence-in-Ultra-High-Dilution-Research.pdf https://doi.org/10.1089/acm.2006.12.437

[85] Demangeat, J.-L. (2013) *Homeopathy*, **102**, 87-105. https://doi.org/10.1016/j.homp.2013.01.001

[86] Van Wassenhoven, M., *et al.* (2017) *Homeopathy*, **106**, 223-239. https://doi.org/10.1016/j.homp.2017.08.001

[87] Konovalov, A.L. and Ryzhkina, L.S. (2014) *Geochemistry International*, **52**, 1207-1226. https://link.springer.com/article/10.1134/S0016702914130072

[88] Bunkin, N.F. and Shkirin, A.V. (2012) *The Journal of Chemical Physics*, **137**, 054707. https://doi.org/10.1063/1.4739530

[89] Häbich, A., *et al.* (2010) *The Journal of Physical Chemistry B*, **114**, 6962-6967. https://doi.org/10.1021/jp911868j

[90] Demangeat, J.-L. (2015) *La Revue d'Homéopathie*, **6**, 125-139. https://doi.org/10.1016/j.revhom.2015.10.002

[91] Sedlak, M. and Rak, D. (2013) *The Journal of Physical Chemistry B*, **117**, 2495-2504. https://doi.org/10.1021/jp4002093

[92] Elia, V., *et al.* (2013) *Water*, **5**, 16-26. http://www.waterjournal.org/volume-5/de-ninno

[93] Elia, V., Yinon, T.A., *et al.* (2016) *Water*, **8**, 1-23.

[94] Ogle, W. (1870) *Medico-Chirurgical Transactions*, **53**, 263-290. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2150467/ https://doi.org/10.1177/095952877005300113

[95] Dyson, G.M. (1938) *Journal of Chemical Technology & Biotechnology*, **57**, 647-651. https://doi.org/10.1002/jctb.5000572802

[96] Turin, L. (1996) *Chemical Senses*, **21**, 773-791. https://academic.oup.com/chemse/article-pdf/21/6/773/1320021/21-6-773.pdf https://doi.org/10.1093/chemse/21.6.773

[97] Turin, L. (2007) The Secret of Scent: Adventures in Perfume and Science of Smell. Harper Collins, 224 p.

[98] Welcome to the Zwiebel Lab. (2012) Vanderbilt University, Ant Chemosensory Biology. http://www.vanderbilt.edu/zwiebellab/Home.html

[99] Benveniste, J. and Guillonnet, D. (2001) Method, System and Device for Producing Signals from a Substance Biological and/or Chemical Activity. https://patents.google.com/patent/US6541978

[100] Mahata, C.R. (2018) *Journal of Homoeopathy University*, **3**, 1-8.

[101] Oliveira, H.M. and Melo, L.V. (2015) Huygens Synchronization of Two Clocks. Scientific Reports 5, Article 11548. https://www.nature.com/articles/srep11548 https://doi.org/10.1038/srep11548

[102] Meessen, A. (2010) Ball Lightning: Bubbles of Elecronic Plasma Oscillations. *International Symposium on Ball Lightning*, Kaliningrad, 21-27 June 2010. http://www.meessen.net/AMeessen/BL-Theory.pdf

[103] https://www.mddionline.com/new-spin-magnetic-sensors-medical-devices

[104] Chattopadhyay, R. and Mahata, C.R. (2016) *International Journal of Complementary & Alternative Medicine*, **3**, 00078 (4 p).

[105] Chattopadhyay, R. and Mahata, C.R. (2016) *International Journal of High Dilution Research*, **15**, 11-17. http://highdilution.org/index.php/ijhdr/article/view/818

[106] Meessen, A. (2017) *Journal of Modern Physics*, **8**, 35-56. http://file.scirp.org/pdf/JMP_2017011015364095.pdf https://doi.org/10.4236/jmp.2017.81004

[107] Meessen, A. (2017) *Journal of Modern Physics*, **8**, 251-267. http://file.scirp.org/pdf/JMP_2017022811200542.pdf https://doi.org/10.4236/jmp.2017.82017

[108] Li, W., *et al*. (1998) Dielectric Spectral Study in Low Frequency of IE Water and Some Other Solvents. Ref. [50], 175-189.

Appendix: The Size of Water Pearls

The basic ideas were presented in Section 2.8. Since $\phi(r \pm \Delta r) = \phi(r) \pm \Delta r \phi'(r)$, where the prime indicates derivation with respect to r, the potential in the external medium ($r \geq R$) is

$$\phi_e(r,\theta) = -P\left(\frac{1}{r^2} + \frac{1}{r\lambda_o}\right)\cos\theta \, e^{-r/\lambda_o} \quad \text{while} \quad \phi_i(r,\theta) = E_i r \cos\theta$$

is the potential in the internal medium ($r \leq R$). P is the total dipole moment of the WP and E_i the electric field inside this nanoparticle. In the external medium it is

$$E_e(r,\theta) = -\phi_e'(r,\theta) = P\left(\frac{2}{r^3} + \frac{2}{r^2\lambda_o} + \frac{1}{r\lambda_o^2}\right)\cos\theta \, e^{-r/\lambda_o}$$

The radius R is then determined by the boundary conditions:

$$\varepsilon_e E_e(R,\theta) + \varepsilon_i E_i(R,\theta) = \sigma(\theta) \quad \text{and} \quad \phi_e(R,\theta) = \phi_i(R,\theta) + w(\theta)$$

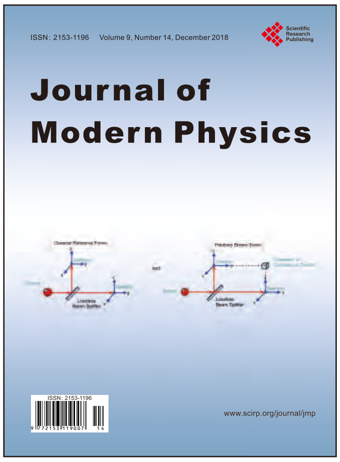
The first relation results from Gauss's law, since the electric fields point away from the interface, which carries a surface *charge* density $\sigma(\theta) = \sigma_o \cos\theta$. The second relation accounts for the *dipole* density $w(\theta)$. It results from the inner and external surface charge densities $\sigma_i \cos\theta$ and $-\sigma_e \cos\theta$. They are separated by a distance $\Delta R \ll R$, but $\sigma_e = \varkappa \sigma_i$ and $\sigma_o = \sigma_i - \sigma_e$, while $w(\theta) = \sigma_e \Delta R \cos\theta$ is negligible. Moreover, $\varepsilon_i = 0$. The boundary conditions yield then the following relations:

$$E_i = -P\left(\frac{1}{R^3} + \frac{1}{R^2\lambda_o}\right)e^{-R/\lambda_o} \quad \text{and} \quad \left(\frac{2}{R^3} + \frac{2}{R^2\lambda_o} + \frac{1}{R\lambda_o^2}\right)e^{-R/\lambda_o} = \frac{\sigma_o}{\varepsilon_e P}$$

We know that $P = \underline{N} q 2a$, where $\underline{N} = (4\pi R^3/3)/d^3$ and $\sigma_i = 2q/d^2$. Thus, $\sigma_o/\varepsilon_e P = (1-\varkappa)(3d/4\pi\varepsilon_e a R^3)$ and

$$\left(2 + 2\eta + \eta^2\right)e^{-\eta} = 0.015(1-\varkappa) \quad \text{where} \quad \eta = R/\lambda_o$$

We ignore the value of $\varkappa = \sigma_e/\sigma_i$, but it results from the last equation that η is nearly constant. Indeed, $\eta = 9.5$ or 10.5 when $\varkappa = 0.5$ or 0.8. Even when $\varkappa = 0.2$, we get $\eta = 9.1$ and $\eta = 11.5$, when $\varkappa = 0.9$. We conclude that $R \approx 10\lambda_o$.



Call for Papers

Journal of Modern Physics

ISSN: 2153-1196 (Print)   ISSN: 2153-120X (Online)
http://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

http://www.scirp.org/journal/jmp **E-mail: jmp@scirp.org**