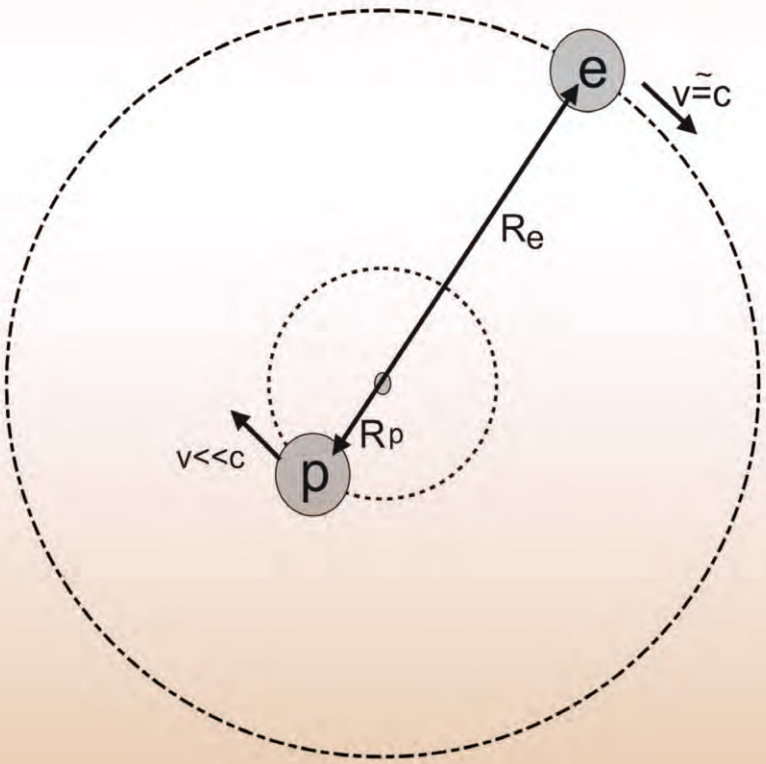
ISSN: 2153-1196 Volume 10, Number 13, November 2019

Scientific Research Publishing

Journal of Modern Physics



ISSN: 2153-1196
9 772153 119007 13

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

Journal Editorial Board

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

Editor-in-Chief

Editorial Board

Journal of Modern Physics, 2019, 10, 1487-1614
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Table of Contents

Volume 10 Number 13 November 2019

The figure on the front cover is from the article published in Journal of Modern Physics, 2019, Vol. 10, No. 13, pp. 1487-1497 by Boris V. Vasiliev.

Journal of Modern Physics (JMP)
Journal Information

SUBSCRIPTIONS

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

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

SERVICES

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

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

COPYRIGHT

PRODUCTION INFORMATION

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

Journal of Modern Physics, 2019, 10, 1487-1497
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

On Nature of Hyperons

Boris V. Vasiliev

Dubna, Russia
Email: bv.vasiliev@yandex.com

How to cite this paper: Vasiliev, B.V. (2019) On Nature of Hyperons. *Journal of Modern Physics*, **10**, 1487-1497.
https://doi.org/10.4236/jmp.2019.1013098

Received: October 8, 2019
Accepted: November 3, 2019
Published: November 6, 2019

Abstract

The purpose of this article is to show that a neutron can have excited states. The well known characteristic feature of the Bohr atom is that its electron shell can exist in a stable ground state or in various excited states. These states differ by integer numbers of de Broglie waves filled in their electronic orbits. Considering neutron to be an analog of the Bohn atom [1] differing in relativistic nature of its electron, a question arises on a possibility for neutron to have similar excited states. The calculations of the properties of these states show that two hyperons Λ^0 and Σ^0 which are usually considered as elementary particles, are excited states of neutron.

Keywords

Independent Researcher, Dubna, Russia

1. Introduction

The foundation of modern physics was laid during the Middle Ages. Since then, the most important achievements include the postulate or principle of W. Gilbert [2]. This postulate formed the basis of all modern natural sciences and created the basis for the successful development of modern physics.

According to this postulate, all theoretical constructs that claim to be scientific must be verified and confirmed experimentally.

It can be formulated in another way: in theoretical physics, all objects must correspond to experimental data, and, even more important, there cannot be objects whose physical properties are principally immensurable experimentally.

For religious people, the existence of angels seems quite normal and natural.

Similar constructions are unacceptable for theoretical physics.

Based on the Gilbert postulate, modern scientific society excludes consideration of the objects whose properties are fundamentally immeasurable.

However, the twentieth century left us a legacy of a number of theoretical

constructions that violate this principle [3].

The fallacy of some provisions of particle physics arose provisions the fact that was based on the model of quarks, and the main method of their description was the construction of tables illustrating the quark structure of particles.

The idea that elementary particles consist of quarks is quite attractive and is confirmed by a number of experiments.

On the other hand, construction of particles by quarks with fractional charge is unsatisfactory. Such particles could not be detected experimentally. The confinement model makes them essentially unobservable, which contradicts the Gilbert principle.

Nevertheless, it is still supposed that in order to understand the world of particles, they need to be collected in tables, sorted by the composition of quarks.

At that new immeasurable quarks are introduced into tables to describe new particles: strange quarks, charmed quarks, beauty quarks, which also differ in colors and aroma.

It is important to emphasize that the ability to classify objects of study by constructing some tables of complex structure (for example, decouplets) proves nothing by itself and can not play the role of experimental proof. At least in such a construction it is necessary first to prove the uniqueness of this classification.

Modern quark theory is based on the fundamental quarks of the lower level u and d. They are needed to explain the important property of neutron: its transformation into proton. However, other properties of neutron cannot be explained by fundamental quarks of the lower level.

All this construction is based on the assumption that neutron is an elementary particle. This hypothesis arose at an early stage of the study of atomic nuclei, when the properties of neutron had not yet been studied.

The question of whether the neutron can be considered a fundamental particle was discussed in the physical community repeatedly in the last century and was solved without relying on measurement data.

One of the first attempts to consider the neutron as a composite particle constructed from proton and electron was made by I. E. Tamm [4]. However, this attempt failed for the reason that became obvious now, it is impossible to construct neutron from proton and a nonrelativistic electron.

In order for theoretical consideration to explain the formation of a composite corpuscle possessing the properties of neutron, it is necessary to consider the unification of a proton with a relativistic electron [5] [6].

This model allows calculating with high accuracy all the main parameters characterizing neutron: its magnetic moment, mass and spin. The mechanism of neutron decay does not require a complicated explanation, but the model allows calculating the energy of this decay.

In addition, this approach makes it possible to explain the nature of nuclear forces on the basis of standard quantum mechanics, whereas gluons, mesons and the strong interaction are excluded from consideration (at least for light nuclei) [1] [6].

Additionally, this model predicts the existence of excited states of neutron.

The existence of excited states of the electron shell is a characteristic feature of the Bohr atom model. In describing the excited states of the electron shell of atom, it is assumed that the degree of excitation is determined by how many de Broglie waves of electron fit on the circumference of the electron orbit.

Using the same principle of formation of excited states, it is possible to determine them for neutron. It turns out that among particles currently classified as elementary, there are those that are not, since their parameters correspond to the excited states of neutron.

Let us consider this question in more detail.

2. The Energy of Interaction of Relativistic Electron with Proton

Consider a composite particle in which an electron having a rest mass m_e and a charge $-e$ is moving around a proton in a circle of radius R_e with a speed $v_e \to c$ (Figure 1).

Since we initially assume that the motion of the electron is likely to be relativistic, it is necessary to take into account the relativistic effect of the growth of its mass:

$$m_e^* = \gamma m_e, \tag{1}$$

where the relativistic factor

$$\gamma = \frac{1}{\sqrt{1-\beta_e^2}} \tag{2}$$

and $\beta_e = \frac{v_e}{c}$.

The rotation of the heavy electron m_e^* does not allow considering the proton as at rest. The proton will also move, revolving around the center of mass common with the heavy electron.

Let's introduce a parameter characterizing the ratio of the mass of a relativistic electron to the mass of proton:

$$\vartheta = \frac{\gamma m_e}{M_p \Big/ \sqrt{1-\beta_p^2}}. \tag{3}$$

Since the ratio of orbit radii is inverse to the ratio of particle masses we get

$$\frac{R_p}{R_e} = \vartheta \tag{4}$$

and radii of orbits of the electron and proton can be written as:

$$R_e = \frac{R_{ep}}{1+\vartheta}, \quad R_p = \frac{R_{ep}\vartheta}{1+\vartheta}. \tag{5}$$

where $R_{ep} = R_e + R_p$.

The relativistic factor characterizing the electron in this case is equal to

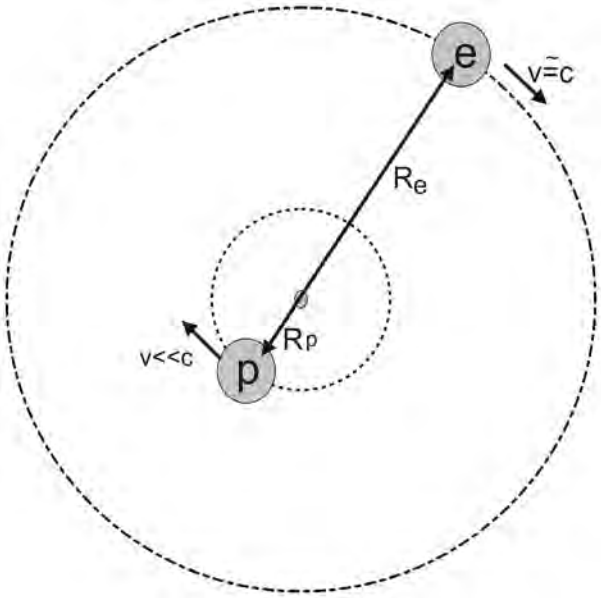


Figure 1. A system consisting of a proton and a heavy (relativistic) electron, revolving around a common center of mass.

$$\gamma = \frac{\vartheta}{\sqrt{1-\vartheta^2}} \frac{M_p}{m_e}. \tag{6}$$

In accordance with Larmor theorem [7], a rotating proton is affected by magnetic field. The magnitude of this field is determined by proton gyromagnetic ratio. The influence of this field will cause the magnetic moment of proton to be oriented perpendicular to the plane of rotation. In other words, due to the interaction with this field, the electron must rotate in the plane of the "equator" of the proton.

2.1. Quantization of Equilibrium Orbit

It can be assumed that, as in the formation of a stable orbit in a hydrogen atom, the orbit of a relativistic electron will be stable if an integer number of de Broglie wavelengths λ_{dB} fits on the circumference of the electron ring $2\pi R_e$, that is:

$$2\pi R_e = n\lambda_{dB}. \tag{7}$$

where n is integer number
and

$$\lambda_{dB} = \frac{2\pi\hbar}{\gamma m_e c}. \tag{8}$$

That is, in accordance with this assumption, the stability condition of the electronic orbit takes the form:

$$\frac{r_c}{R_e} = \frac{\vartheta}{n\sqrt{1-\vartheta^2}} \frac{M_p}{m_e} = \frac{\gamma}{n} \tag{9}$$

where $r_c = \frac{\hbar}{m_e c}$ is Compton radius.

2.2. The Kinetic Energy of the System of Relativistic Electron + Proton

The kinetic energy of a relativistic electron is expressed by the equality:

$$\mathcal{E}_{kin}^{e} = (\gamma - 1) \cdot m_e c^2 \tag{10}$$

Due to the assumption of the electron to be ultrarelativistic

$$\mathcal{E}_{kin}^{e} \approx \gamma \cdot m_e c^2 \tag{11}$$

In this case, the centrifugal force acts on the electron:

$$\mathcal{F}_1 = \gamma m_e \left[\omega \left[\omega, R_e\right]\right] = \frac{\gamma m_e c^2}{R_e} \tag{12}$$

The kinetic energy of the proton is equal to:

$$\mathcal{E}_{kin}^{p} = \left(\frac{1}{\sqrt{1-\vartheta^2}} - 1\right) \cdot M_p c^2 \tag{13}$$

2.3. The Coulomb Interaction in the System of Relativistic Electron + Proton

The energy of Coulomb attraction between a proton and a relativistic electron is proportional to the relativistic factor γ [7], §24:

$$\mathcal{E}_C = -\gamma \frac{e^2}{R_{ep}} = -\gamma \frac{\alpha r_c}{R_e (1+\vartheta)} m_e c^2. \tag{14}$$

where $\alpha = \frac{e^2}{\hbar c}$ is the fine structure constant.

Therefore, the Coulomb attraction force acting between these particles is equal to

$$\mathcal{F}_2 = -\gamma \frac{e^2}{R_{ep}^2} = -\gamma \frac{\alpha}{(1+\vartheta)^2} \frac{r_c}{R_e} \frac{m_e c^2}{R_e}. \tag{15}$$

2.4. The Magnetic Interaction of a Rotating Relativistic Electron

2.4.1. Magnetic Energy of the Electron Current Ring

An additional contribution to the kinetic energy of the system is made by the magnetic energy of a rotating electron.

The energy of the magnetic field created by the rotation of electron tends to break the ring of electron current. This energy depends on the magnitude of the magnetic flux in the ring Φ and the current J which creates it:

$$\mathcal{E}_{\Phi} = \frac{\Phi J}{2}. \tag{16}$$

Due to the fact that the electron orbit is quantized, the magnetic flux penetrating the ring of radius R_e should be equal to the magnetic flux quantum

Φ_0

$$\Phi = \Phi_0 = \frac{2\pi\hbar c}{e}. \tag{17}$$

By definition the magnetic flux in the ring is determined by the current J_0 and the area of the ring S_0:

$$\mu_0 = J_0 \cdot S_0 \tag{18}$$

i.e.

$$\mathcal{E}_{\Phi_e} = \frac{e^2}{R_e}\frac{1}{2\alpha}\frac{r}{R_e} = \frac{1}{2n}\frac{\vartheta}{\sqrt{1-\vartheta^2}}\cdot M_p c^2. \tag{19}$$

The force arising at the same time, tending to break the current ring, turns out to be equal

$$\mathcal{F}_3 = \frac{\gamma}{2n}\frac{m_e c^2}{R_e}. \tag{20}$$

The magnetic energy created by the rotation of a proton is much less:

$$\mathcal{E}_{\Phi_p} = \frac{\sqrt{2}\cdot\vartheta^2}{\sqrt{1-\vartheta^2}}\cdot M_p c^2. \tag{21}$$

The force corresponding to this energy is applied to proton and does not directly affect the electron equilibrium orbit.

2.4.2. Interaction of Electron with Magnetic Field of Proton

In the present case a proton possesses two magnetic moments. This is its own internal magnetic moment:

$$\mu_p = \frac{\xi e\hbar}{2M_p c} \tag{22}$$

and the orbital magnetic moment which occurs due to the fact that proton rotates in an orbit of radius R_p:

$$\mu_{0p} = \frac{e\vartheta R_p}{2} \tag{23}$$

Therefore, the energy of interaction of rotating electron with the proton magnetic field consists from two components:

$$\mathcal{E}_\mu = \pm\frac{\gamma e}{2R_e^2}\left(\mu_{0_p} - \mu_p\right). \tag{24}$$

In order for the system energy to be less, the magnetic moments μ_p and μ_{0_p} must be oppositely directed. But the total contribution of the energy of this interaction can be either positive or negative. It depends on the direction of electron rotation relative to the orientation of the proton magnetic moment. Therefore, in the future, when solving these equations, it will be necessary to take into account both options with different signs.

The force that acts on the rotating electron can be written as:

$$\begin{aligned}\mathcal{F}_4 &= \pm\gamma e\beta\left(\frac{\mu_{0_p}}{R_e^3}-\frac{\mu_p}{R_{ep}^3}\right)=\pm\gamma e\left(\frac{\mu_{0_p}}{R_e^3}-\frac{\mu_p}{R_e^3(1+\vartheta)^3}\right)\\ &=\pm\gamma\frac{m_e c^2}{R_e}\left(\frac{\vartheta^2}{2}-\frac{\xi}{(1+\vartheta)^3}\frac{\vartheta}{2n\sqrt{1-\vartheta^2}}\right)\frac{\vartheta}{2n\sqrt{1-\vartheta^2}}\alpha\frac{M_p}{m_e}.\end{aligned} \tag{25}$$

where $\xi \approx 2.79$ is the proton magnetic moment expressed in Bohr magnetons.

The magnetic moment of electron is not considered because, as will be shown below, the generalized momentum (spin) of the electron orbit is equal to zero and there is no direction for the selected orientation of the electron magnetic moment in the system.

3. Equilibrium Electron Orbit

The equilibrium condition for the electron orbit is:

$$\sum_{i=1}^{4}\mathcal{F}_i = 0. \tag{26}$$

At summing of Equation (12), Equation (15), Equation (20) and Equation (25) after simplifying transformations taking into account Equation (9) we get:

$$1+\frac{1}{2n}-\left(\frac{\vartheta}{n\sqrt{1-\vartheta^2}}\frac{\alpha M_p}{m_e}\right)\left[\frac{1}{(1+\vartheta)^2}\pm\left(\frac{\vartheta^2}{2}-\frac{\xi}{2n(1+\vartheta)^3}\frac{\vartheta}{\sqrt{1-\vartheta^2}}\right)\right]=0. \tag{27}$$

The double sign ± before the last term on the left side of this equality is explained by the fact that the direction of this force depends on the direction of rotation of relative to the magnetic field created by the magnetic moment of proton.

To find the electron orbit with minimum energy, the solution of this equation with respect to ϑ must be carried out for each directions of the electron rotation.

3.1. The State with $n = 1$

Under this condition, one needs to find a solution to the equation:

$$\begin{aligned}&1+\frac{1}{2}-\left(\frac{\vartheta}{\sqrt{1-\vartheta^2}}\frac{\alpha M_p}{m_e}\right)\left[\frac{1}{(1+\vartheta)^2}\right]\\ &-\left(\frac{\vartheta}{\sqrt{1-\vartheta^2}}\frac{\alpha M_p}{m_e}\right)\left[\frac{\vartheta^2}{2}-\frac{\xi}{2(1+\vartheta)^3}\frac{\vartheta}{\sqrt{1-\vartheta^2}}\right]=0.\end{aligned} \tag{28}$$

As a result, the solution of this equation is

$$\vartheta = 0.1991. \tag{29}$$

3.2. The State with $n = 2$

Under this condition the equation is

$$1+\frac{1}{2\cdot 2}-\left(\frac{\vartheta}{2\sqrt{1-\vartheta^2}}\frac{\alpha M_p}{m_e}\right)\left[\frac{1}{(1+\vartheta)^2}\right]$$
$$+\left(\frac{\vartheta}{2\sqrt{1-\vartheta^2}}\frac{\alpha M_p}{m_e}\right)\left[\frac{\vartheta^2}{2}-\frac{\xi}{2\cdot 2(1+\vartheta)^3}\frac{\vartheta}{\sqrt{1-\vartheta^2}}\right]=0. \tag{30}$$

The solution to this equation is

$$\vartheta = 0.263. \tag{31}$$

3.3. The State with $n = 3$

At that the equation is

$$1+\frac{1}{2\cdot 3}-\left(\frac{\vartheta}{3\sqrt{1-\vartheta^2}}\frac{\alpha M_p}{m_e}\right)\left[\frac{1}{(1+\vartheta)^2}\right]$$
$$-\left(\frac{\vartheta}{3\sqrt{1-\vartheta^2}}\frac{\alpha M_p}{m_e}\right)\left[\frac{\vartheta^2}{2}-\frac{\xi}{2\cdot 3(1+\vartheta)^3}\frac{\vartheta}{\sqrt{1-\vartheta^2}}\right]=0 \tag{32}$$

and its solution is

$$\vartheta = 0.479. \tag{33}$$

4. The Particle Magnetic Moment

The particle magnetic moment is the sum of the proton magnetic moment and magnetic moments of orbital currents of electron and proton.

The total magnetic moment generated by of both circular currents

$$\mu_0 = -\frac{e\beta_e R_e}{2}+\frac{e\beta_p R_p}{2}=\frac{eR_{ep}}{2}\frac{(1-\vartheta^2)}{(1+\vartheta)}=\frac{eR_{ep}}{2}(1-\vartheta). \tag{34}$$

If to express this moment in the magnetons of Bohr μ_B, we get

$$\xi_0 = \frac{\mu_0}{\mu_B} = -\frac{(1-\vartheta^2)\sqrt{1-\vartheta^2}}{\vartheta^2}. \tag{35}$$

Thus, the magnetic moment of the electron orbit:

$$\mu_0 = \left[-\frac{(1-\vartheta^2)\sqrt{1-\vartheta^2}}{\vartheta}\right]. \tag{36}$$

Summing it with the proton magnetic moment, we get

$$\mu_{total} = \left[-\frac{(1-\vartheta^2)\sqrt{1-\vartheta^2}}{\vartheta}+2.79\right]. \tag{37}$$

These values at different ϑ are shown in **Table 1**.

It should be noted that the magnetic moment of Σ^0-hyperon in [8] is designated as the transition moment of $\mu^0_{\Sigma\Lambda}$.

5. Mass of Particles

The mass of a composite particle is determined by the sum of the rest masses of

Table 1. Comparison of calculated values of magnetic moments with measurement data.

n	ϑ	μ_0	μ_{total}	experimental	Ref.
		Equation (36)	Equation (37)	data	
$n=1$	0.1991	−4.727	−1.9367	$\mu_{n_0} = -1.9130427 \pm 0.0000005$	[8]
$n=2$	0.263	−3.4147	−0.6247	$\mu_{\Lambda^0} = -0.613 \pm 0.004$	[8]
$n=3$	0.479	−1.4121	1.3779	$\mu_{\Sigma^0_{\Sigma\Lambda}} = 1.61 \pm 0.08$	[8]

the particles, their relativistic kinetic energy and the mass defect arising from the potential energy of their internal interaction. Calculate these contributions.

5.1. Kinetic Energy of Electron and Proton

Summing Equations (11), (13), (19) and (21) we obtain

$$\mathcal{E}(kin) = \frac{\vartheta}{\sqrt{1-\vartheta^2}}\left[1+\left(\frac{1}{\sqrt{1-\vartheta^2}}-1\right)\frac{\sqrt{1-\vartheta^2}}{\vartheta}+\left(\frac{1}{2n}+\sqrt{2}\vartheta\right)\right]\cdot M_p c^2 \tag{38}$$

5.2. Potential Energy of Electron and Proton

Summing Equations (14) and (24) we obtain

$$\mathcal{E}(pot) = \frac{\alpha M_p}{nm_e}\left[\frac{1}{1+\vartheta}-\frac{\vartheta^2}{2}\left(1-\frac{1}{(1+\vartheta)^3}\cdot\frac{\xi}{n\cdot\vartheta\sqrt{1-\vartheta^2}}\right)\right]\left(\frac{\vartheta}{\sqrt{1-\vartheta^2}}\right)^2\cdot M_p c^2. \tag{39}$$

5.3. Neutron and Hyperon Masses

The total mass of proton and electron at taking in to account their energies:

$$\begin{aligned} M_{total} &= m_e + M_p + \frac{\mathcal{E}(kin)}{c^2} - \frac{\mathcal{E}(pot)}{c^2} \\ &= m_e + M_p + \frac{\vartheta}{\sqrt{1-\vartheta^2}}\left[1+\left(\frac{1}{\sqrt{1-\vartheta^2}}-1\right)\frac{\sqrt{1-\vartheta^2}}{\vartheta}+\left(\frac{1}{2n}+\sqrt{2}\vartheta\right)\right]\cdot M_p \\ &\quad -\frac{\alpha M_p}{nm_e}\left[\frac{1}{1+\vartheta}-\frac{\vartheta^2}{2}\left(1-\frac{1}{(1+\vartheta)^3}\cdot\frac{\xi}{n\cdot\vartheta\sqrt{1-\vartheta^2}}\right)\right]\left(\frac{\vartheta}{\sqrt{1-\vartheta^2}}\right)^2\cdot M_p \end{aligned} \tag{40}$$

This formula allows us to calculate masses of particles in question as a function of the parameter n. The results of calculations are summarized in **Table 2**.

The sum of kinetic and potential energy thus obtained must correspond to the energy released during the decay of the particle. For the neutron, this estimate is in qualitative agreement with the measured data.

6. Spin Particles

Since in the relativistic case the vector-potential takes the form [7], §24:

$$A = \gamma\left(A' + \beta\varphi'\right), \tag{41}$$

Table 2. The comparison of calculated particle mass values with measurement data.

n	$\frac{\mathcal{E}_{kin}}{c^2}$	$\frac{\mathcal{E}_{pot}}{c^2}$	M_{total}	experimental	$\Delta = \frac{M_{exp} - M_{calc}}{M_{exp}}$
			Equation (40)	data	
$n = 1$	$702m_e$	$700m_e$	$1839m_e$	$M_{n_0} = 1837m_e$	0.001
$n = 2$	$879m_e$	$778m_e$	$1938m_e$	$M_{\Lambda^0} = 2183m_e$	0.11
$n = 3$	$2103m_e$	$1740m_e$	$2200m_e$	$M_{\Sigma^0} = 2335m_e$	0.06

the force that acts on the charge of a relativistically rapidly rotating particle can be represented as:

$$F_e = \gamma e \cdot rotA, \tag{42}$$

and as a result, taking into account the Equation (37) to obtain a condition for the generalized momentum of the particle

$$P_0 = \gamma mc + \gamma \frac{e}{c} A = 0. \tag{43}$$

Thus, in the case under consideration, the moment of the generalized momentum of rotating particles

$$S_0 = \left[R_e, P_0\right] = 0. \tag{44}$$

For this reason, the total spin of the particles in question is 1/2 because it is created by the spin of the proton.

A detailed computation of neutron spin is considered in [1] [6].

7. Conclusions

It should be emphasized that the above estimates of the basic parameters of the corpuscles under consideration are obtained in a simple and usual way. They do not contain any hidden fitting parameters. The agreement that the calculated parameters show when compared with the corresponding measured values leads to important conclusions.

As a consequence, neutron, Λ^0- and Σ^0-hyperons (as well as π-mesons and μ-mesons [9]), cannot be considered elementary particles, as it is commonly thought at present.

There is no need to introduce strange quarks to describe Λ^0- and Σ^0-hyperons (just as there is no need to introduce basic u and d quarks to describe neutron decay [1] [6]).

The exclusion of hyperons, as well as mesons and neutrons, from the table of elementary particles deconstructs these tables, built on the hypothesis of the existence of quarks with a fractional charge, thus destroys the hypothesis of the existence of the baryon decuplet because it included Σ^0-hyperons.

It can be assumed that many other particles like Σ^+ and Σ^- are also not elementary particles, but are short-living excited states of other constituent corpuscules.

Conflicts of Interest

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

References

[1] Vasiliev, B.V. (2017) *SciFed Journal of Nuclear Science*, **1**, 1. http://scifedpublishers.com/fulltext/the-electromagnetic-model-of-neutron-and-nature-of-nuclear-forces/21910

[2] Gilbert, W. (1600) De magneto magneticisque corporibus et de magno magnete tellure, London.

[3] Vasiliev, B.V. (2018) *Journal of Modern Physics*, **9**, 2101-2124. https://www.scirp.org/Journal/PaperInformation.aspx?PaperID=87652 https://doi.org/10.4236/jmp.2018.912132

[4] Tamm, I.E. (1934) *Nature*, **134**, 1010-1011. https://doi.org/10.1038/1341010c0

[5] Vasiliev, B.V. (2016) *International Journal of Modern Physics and Application*, **3**, 25-38. http://www.aascit.org/journal/archive2?journalId=909&paperId=3935

[6] Vasiliev, B.V. (2015) *Journal of Modern Physics*, **6**, 648-659. https://doi.org/10.4236/jmp.2015.65071 http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=55921 http://n-t.ru/tp/ns/op.htm

[7] Landau, L.D. and Lifshitz, E.M. (1971) The Classical Theory of Fields. In: *Volume* 2 *of A Course of Theoretical Physics*, Pergamon Press, NY, §24.

[8] Tanabashi, M., *et al.* (2018) *Physics Reviews D*, **98**. 030001 https://doi.org/10.1103/PhysRevD.98.030001

[9] Vasiliev, B.V. (2019) *Journal of Modern Physics*, **10**, 1-7. http://www.scirp.org/pdf/JMP_2019011014591744.pdf https://doi.org/10.4236/jmp.2019.101001

Journal of Modern Physics, 2019, 10, 1498-1504
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Informational Entropy as a Source of Life's Origin

Shahram Yazdani

Department of Pediatrics, David Geffen School of Medicine at University of California, Los Angeles, CA, USA
Email: Syazdani@mednet.ucla.edu

How to cite this paper: Yazdani, S. (2019) Informational Entropy as a Source of Life's Origin. *Journal of Modern Physics*, **10**, 1498-1504.
https://doi.org/10.4236/jmp.2019.1013099

Received: October 16, 2019
Accepted: November 10, 2019
Published: November 13, 2019

Abstract

While the second law of thermodynamics suggests that our universe is driven by the tendency towards disorder, living organisms seem to exempt themselves by creating physiologic complexity. Since genetic material is life's blueprint, better understanding of the origins of life is predicated on deciphering the conditions that allowed the formation of this complex molecule with its unique properties. In this article, we propose and examine the hypothesis that informational entropy models would allow for the formation of complex organic molecules with genetic properties, without the disruption of the second law of thermodynamics. Therefore, we demonstrate that formation of life's blueprint may have initially been derived by informational entropy by means of decomplexification of the materials with higher informational entropy content, leading to the formation of primitive genetic molecules.

Keywords

Entropy, Informational Entropy, Genetic Material, Second Law of Thermodynamics

1. Introduction

The ability of living organisms to defy the natural tendency of all matter towards decay by creating order and evolving into complex organisms has been the subject of much fascination since the ancient times. Perhaps the earliest implementation of the role of disorder in relation to life was conveyed to us through Greek mythology where Gaea, the earth and mother of all things, sprang from Chaos. During the more modern times, the eighteenth century German chemist and physician Georg Ernst Stahl described life as a *conservatio mixtionis corporis*, against the tendency to decompose. Later, Erwin Schrodinger in his 1944 book titled "What is life", described the tendency of the living organisms to seemingly

exempt themselves from the second law of thermodynamics by avoiding decay through a state of non-equilibrium with their surroundings [1]. Thus, while our universe is driven by the tendency towards higher entropy or disorder, living organisms constantly strive to increase or maintain physiological complexity in order to avoid the increase in their entropy to a state of equilibrium with the surrounding environment. As thermodynamically open systems, living organisms must harness free energy or enthalpy from food and chemical substrate to create an ideally equal balance between the entropy created within the system and the entropy exported [2]. Thus, as Schrodinger had concluded, all living organisms continuously create and export positive entropy while maintaining as he called it, "negative entropy". Furthermore, since the transformation of inorganic compounds to organic compounds was initially demonstrated in 1828 by the German chemist Friedrich Wöhler, who artificially obtained urea by treating silver cyanate with ammonium chloride, the transformation of inanimate being to life form is no longer hindered by the absence of preceding organic compounds. Thus, a better understanding of the forces that may resolve the classical thermodynamic dilemmas of life forms, paves the way for peering into possible conditions and requirements that may have initially led to the formation of life, or at the very least, its most fundamental component, life's genetic blueprint (e.g. DNA, RNA).

Yet, formation of life's molecular blueprint and its surrounding structures from the less complex substrates requires temporary violation of the second law of thermodynamics by reducing the entropy of the substrates through their organization into more complex compounds and eventually a living organism with the promise of increased consumption of food material, export of entropy, and thus overall increase in the system's entropy. To overcome this challenge, Horowitz and England have proposed an in silico model of molecules where inanimate material settings may act "life-like" by maintaining a "far-from-equilibrium" steady state [3].

In this article, we demonstrate that information has physical properties that provide an essential component of the thermodynamics of living organisms and can provide a solution in regards to the thermodynamic pathway that paved the way to de novo formation of life from inanimate organic or inorganic substrates. In other words, we demonstrate that the initial formation of the primitive genetic material, which is one of the most fundamental components of life, is not formed through a series of complex physical reactions that are only applicable in theoretical timeframes or environments, but in fact is a natural phenomenon that can be explained through the observed reactions within the physical environment.

Therefore, we propose that,

1) Shannon's informational entropy in the context of biology, suggests that information is an essential deriving factor in the formation of life.

2) Inanimate inorganic compounds can not only provide a substrate for life, but also provide informational entropy content for the earliest forms of genetic

material, thus providing a solution for the thermodynamic derive that led to the formation of life.

2. Information and Life

Life's defiance of the second law of thermodynamics within its system has led to the exploration of alternative or complementary physical laws that may explain life's tendency to thrive through the export of entropy to its surrounding in order to maintain disequilibrium between the living organism, and its surrounding environment.

As Sanchez has demonstrated through the "egg conundrum", which calculates the thermodynamics of a fertilized egg compared to an unfertilized one, the statistical and classical thermodynamics cannot explain the formation of life [4]. Furthermore, Sanchez calculates the assignable statistical entropy content of each codon using Gibbs entropy formula.

$$S = -k\sum_{l} Pl \ln(Pl) \tag{1}$$

where k is the Boltzmann's constant, Pl in this setting represents the probability of a unique consecutive "linear" sequence of N codons in the l_{th} state. Using this approach, Sanchez concludes that the negative entropy assignable to each codon is estimated as,

$$S = -3.06 \text{ K} \tag{2}$$

As the equation above demonstrates, the statistical negative entropy content of each codon does not fully account for the positive entropy generated by an organism [4]. Therefore, inclusion of the informational content of DNA is essential to account for its total negative entropy content.

However, while including the informational content of DNA as a source of its negative entropy offers a solution for the inadequacy of the statistical thermodynamic models, it implies that formation of genetic material requires increase in stored information, which further poses the question as to what conditions originally derived formation and accumulation of information in the initial molecular blueprints that would eventually lead to life formation.

3. Substrate for Information

To better understand conditions under which the primitive form of a complex, information-rich molecule such as primitive RNA or DNA may have been initially formed, we propose a look at the informational content of some of the comparatively more complex inorganic compounds found in nature.

We begin by introducing the Shannon information formula

$$H = -\sum_{i=1}^{s} Pi \log_2 Pi \ \left(\text{bits/symbol}\right) \tag{3}$$

where H is the amount of information contained in a system with n symbols of s different types occurring with the probability Pi. The importance of this formula is its universal application to any form of message, regardless of its medium.

Therefore, blueprint of living organisms and inanimate sources share certain commonalities in the rules governing information and its relation to physical properties as described in the example below.

Consider the example of Crystals. As Krivovichev has demonstrated, the structural complexity and configuration of crystal structures add informational entropy to crystal structures that go beyond the molecular content of each crystal [5]. The informational content I_G per atom appearing in a crystal is calculated as,

$$I_G = \sum_{i=1}^{k} Pi \ log_2 Pi \ \left(\text{bits/atom}\right) \tag{4}$$

where k is the number of crystallographic orbits, denoted as i, and Pi is the probability of a randomly chosen atom appearing in the i-th crystallographic orbit.

Using Shannon's information equation, Krivovichev demonstrates that even in a simple crystal with all atoms equivalent and distributed into various asymmetrical crystallographic orbits, the configuration of orbits and arrangement of various crystal "cells" creates informational content that in fact holds physical properties, makes negative contribution to the entropy of the crystal, and therefore, should be accounted for in the total entropy of the crystal.

4. Information as the Common Denominator

As Equations (2) and (4) imply, DNA and crystals carry information through the arrangement of their analogous informational building blocks. More specifically, the configuration of atoms in a crystal carries information, similarly to the linear arrangement of codons in the case of a strand of DNA. Furthermore, as Landauer suggested (1961, 1996), erasure of each bit of information requires energy loss (expressed in degrees Kelvin) and thus, information and energy are interlinked [6] [7].

$$kT \ln 2$$

where K is *Boltzman's constant* and temperature is in Kelvin.

As Landauer suggests, for each bit of information lost at Temperature T, $k\ln 2$ of energy is released into the environment, which in turn contributes to the entropy S of the environment.

This suggests that complexification of a crystal, and formation of any form of life's blueprint should both be considered within the context of energy and information exchange, within their surrounding environment.

Considering the information-rich content of many complex crystals, it is conceivable that conditions that allowed information and thermodynamic exchange between the substrates of the primitive genetic material and comparatively more complex crystals in their surroundings, led to the formation of the earliest genetic material [7]. Therefore, the net result of the simplification of these crystal structures, and hence increase in their informational and thermodynamic entropy, led to complexification and decrease in the thermodynamic and informa-

tional entropy of the substrates of the earliest forms of codons. Hence, this transference of negative informational and thermodynamic entropy from the comparatively more complex molecules such as complex crystals to genetic substrates, eliminates the need for a hypothetical temporary decrease in the entropy of the system with the promise of eventual increase in the entire system's entropy that would arrive after the complete formation of a functioning organism. Furthermore, considering the unknown composition of the initial blueprints, minerals of complex composition may have in fact provided both the substrate and entropy (informational and statistical) of these molecular blueprints. Therefore, we propose that inorganic molecules may have potentially, not only provided the substrate to form organic molecules, as Friedrich Wöhler had demonstrated, but also acted as the source of informational entropy through a series of chemical reactions that led to decomplexification of these inorganic molecules coupled with transference of their informational entropy to the newly formed organic molecules. Thus, this exchange of informational entropy can satisfy the second law of thermodynamics while providing adequate informational entropy to allow these primitive molecules to carry gene-like properties.

However, the proposed hypothesis greatly hinges upon the existence of inorganic molecules (e.g. mineral crystals) with complex structures that would have been geologically formed prior to the formation of the first life form, while containing adequate complexity to contain informational entropy that exceeds the primitive genetic material.

Although the informational content and composition of the earliest genetic material is beyond the scope of this paper and may have significantly varied from the informational content of the modern DNA, a comparison of the more primitive crystals' informational content with the more modern DNA codons provides some clues as to the plausibility of this hypothesis. Therefore, if a unit cell of a crystal holds more information than a comparable DNA codon, then it is possible that decomplexification of an analogous crystal could have led to the formation of a codon of comparatively lesser informational content and complexity, and therefore, the second law of thermodynamics is preserved.

To compare the informational content of a unit of crystal with the entropy of each codon of DNA, we must first consider the energy content of each bit of information as noted in Landauer's equation. The ratio of the energy released and therefore contributed to the entropy of its environment from all or part of the informational content of a crystal in the process of decomplexification as calculated through Landauer's equation, divided by the Gibbs entropy of each codon of DNA must be equal or exceed 1.

$$I_c\left(kT\ln 2\right)/3.06k = I_cT0.227 > 1$$

$$I_cT > 1/0.227 = 4.405$$

where I_c is the informational content of a unit of crystal, k is the Boltzman's constant and T is the temperature of the environment in Kelvin. As this indicates, as long as the product of the temperature and informational content of each unit

cell of a crystal exceeds 4.405, the ratio above is satisfied.

A brief review of some of the structurally simple crystals in the initial geological formations (>4.55 Ga) preceding our estimated time for the formation of primitive life indicates that a variety of crystals had moderate level of informational content that may have reached up to 200 bits/unit cell, far exceeding the informational content needed to form each codon of a strand of DNA [8] [9].

Minerals and their crystal structure also offer a unique solution since despite their low entropy and high informational content, they are inevitably formed and evolve into more complex structures under geological conditions [9]. Therefore, geological conditions that naturally led to the formation of complex inorganic structures such as minerals, may have in turn led to the formation of life and its essential components through a series of complexification (*i.e.* crystal formation), and subsequent decomplexification of crystals that lead to complexification of inorganic and organic molecular building blocks of the initial genetic material.

5. Conclusion

Statistical thermodynamic models alone fail to explain the conditions under which life could have initially been formed. However, integration of the informational entropy models provides a plausible scenario for life formation under known natural conditions. This model suggests that informational and energy exchange between the initial substrates of the primitive genetic material and the more complex molecules and structures may have made life possible. The presence of crystals and minerals with low entropy and high informational content preceding life formation, offers a plausible source of such complex molecules. Further experimentation is needed to test this hypothesis by delineating the potential chemical reactions through which inorganic molecules in the environment surrounding the newly formed organic molecules can exchange and contribute informational entropy, in order to enable gene-like properties that are essential for life formation.

Author Contributions

The sole author of this article has been actively involved in forming this hypothesis and writing the final copy of this manuscript. The author takes full responsibility in the integrity of this study and its content. The author has read the submitted manuscript and is fully aware and in agreement with the accuracy of the presented information and conclusions of this study.

Competing Financial Interest Statement

The author of this article has no financial conflicts of interest in relation to the outcomes of this study. This study has been financed through internal sources.

Conflicts of Interest

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

References

[1] Schrödinger, E. (1944) What Is Life? Cambridge University Press, Cambridge.

[2] Prigogine, I. (1947) Etude thermodynamique des Phenomenes Irreversibles. Desoer Liege.

[3] Horowitz, J.M. and England, J.L. (2017) *Proceedings of the National Academy of Sciences of the United States of America*, **114**, 7565-7570. https://doi.org/10.1073/pnas.1700617114

[4] Sanchez, I.C. (2011) *Journal of Modern Physics*, **7**, 654-657. https://doi.org/10.1073/pnas.1700617114

[5] Krivovichev, S.V. (2016) *Acta Crystallographica*, **72**, 274-276. https://doi.org/10.1107/S205252061501906X

[6] Landauer, R. (1961) *IBM J. Res.*, **5**, 183-191. https://doi.org/10.1147/rd.53.0183

[7] Landauer, R. (1996) *Physics Letters A*, **217**, 188-193. https://doi.org/10.1016/0375-9601(96)00453-7

[8] Krivovichev, S.V. (2013) *Mineralogical Magazine*, **77**, 275-326. https://doi.org/10.1180/minmag.2013.077.3.05

[9] Hazen, R.M., Papineau, D., Bleeker, W., Downs, R.T., Ferry, J.M., McCoy, T.J., Sverjensky, D.A. and Yang, H. (2008) *American Mineralogist*, **93**, 1693-1720. https://doi.org/10.2138/am.2008.2955

Journal of Modern Physics, 2019, 10, 1505-1521
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Design of Hanman Entropy Network from Radial Basis Function Network

M. Hanmandlu[1], Rao Shivansh[2], Shantaram Vasikarla[3]

[1]CSE Department MVSR Engg. College, Nadergul, Hyderabad, India
[2]Pennsylvania State University State College, PA, USA
[3]CSE Department California State University Northridge, CA, USA
Email: mhmandlu@ee.iitd.ernet.in

How to cite this paper: Hanmandlu, M., Shivansh, R. and Vasikarla, S. (2019) Design of Hanman Entropy Network from Radial Basis Function Network. *Journal of Modern Physics*, **10**, 1505-1521.
https://doi.org/10.4236/jmp.2019.1013100

Received: August 25, 2019
Accepted: November 10, 2019
Published: November 13, 2019

Abstract

Different learning algorithms have been developed in the literature for training the radial basis function network (RBFN). In this paper, a new neural network named as Hanman Entropy Network (HEN) is developed from RBFN based on the Information set theory that deals with the representation of possibilistic uncertainty in the attribute/property values termed as information source values. The parameters of both HEN and RBFN are learned using a new learning algorithm called JAYA that solves the constrained and unconstrained optimization problems and is bereft of algorithm-specific parameters. The performance of HEN is shown to be superior to that of RBFN on four datasets. The advantage of HEN is that it can use both information source values and their membership values in several ways whereas RBFN uses only the membership function values.

Keywords

RBFN, HEN, Gradient Descent (GD), Pseudo-Inverse, JAYA

1. Introduction

The artificial neural networks (ANNs) that include back propogation (BP) networks [1], radial basis function networks (RBFNs) [2], counter propagation networks [3] to mention a few show their power in data classification, pattern recognition and function approximation. In this paper, we are mainly concerned with incorporating a new learning agorithm, called JAYA into the architecture of RBFN to mitigate the drawbacks of its gradient descent learning.

A radial basis function network (RBFN) [4] [5] is a three-layer feed-forward neural network. Each hidden layer neuron evaluates its kernel function on the

incoming input. The network output is simply a weighted sum of the values of the kernel functions in the hidden layer neurons. The value of a kernel function is highest when the input falls on its center and decreases monotonically as it moves away from the center. A Gaussian function is normally used as the kernel function. The training of an RBFN is done by finding the centers and the widths of the kernel functions and the weights connecting the hidden layer neurons to the output layer neurons.

Next, we will foray into the learning domain. Finding the global optimum of a function is the main task of many of the scientific applications. Gradient descent approach is widely used but it suffers from local minima. Another limitation is that it cannot be used in optimization problems that have non-differentiable objective functions. Many modern population based heuristic algorithms focus on finding a near optimum solution to overcome this requirement of differentiability associated with gradient descent learning.

A brief survey of the population based heuristic algorithms will enlighten the readers how much work has been done in the domain of learning. These algorithms can be clubbed into two important groups: evolutionary algorithms (EA) and swarm intelligence (SI) based algorithms. Some of the recognized evolutionary algorithms are: Genetic Algorithm (GA), Evolution Strategy (ES), Evolution Programming (EP), Differential Evolution (DE), Bacterial Foraging Optimization (BFO), Artificial Immune Algorithm (AIA), etc. Some of the well known swarm intelligence based algorithms are: Particle Swarm Optimization (PSO), Shuffled Frog Leaping (SFL), Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Fire Fly (FF) algorithm, etc. Besides the evolutionary and swarm intelligence based algorithms, there are some other algorithms that work on the principles of different natural phenomena. Some of them are: Harmony Search (HS) algorithm, Gravitational Search Algorithm (GSA), Biogeography-Based Optimization (BBO), Grenade Explosion Method (GEM), etc. All the evolutionary and swarm intelligence based algorithms are probabilistic algorithms that require common controlling parameters like population size, number of generations, elite size, etc. Besides the common control parameters, different algorithms require their own algorithm-specific control parameters. A recent meta-heuristic learning method called Human Effort for Achieving Goals (HEFAG) by Jyotsana and Hanmandlu contains the comparison of several learning methods in [6]. A new learning algorithm called JAYA is developed in [7] to overcome the need for the algorithm-specific parameters but the need for the common control parameters still exists. This algorithm helps the initial solutions move towards the best solution by avoiding the worst solution.

2. Design of RBFN

For the detailed study on artificial neural networks (ANN) and fuzzy systems and their applications readers may refer to Jang *et al.* [8]. As Multilayer Perceptron (MLP) is a major leap in ANNs and RBFN has arisen out of

simplying computational burden involved in MLP; hence it is widely used [9] for the traditional classification problems. A comparison between the traditional neural networks and RBFN is presented in [10].

RBFN deals with attrubute/feature values that are clustered. The attribue values in a cluster are fitted with the radial basis function which is another name for Gaussian function. RBFN fuzzifies the attribute values in a cluster into the membership function values. Each RBFN neuron stores a cluster centre or centroid, which is initially taken to be one of the samples from the training set. When we want to classify a new input, each neuron computes the Euclidean distance between the input and its centroid and computes the membership function using the standard deviation or width of the Gaussian function. The output of the RBFN is a weighted sum of the membership function values as shown in **Figure 1**. In this μ_i denotes the i^{th} membership function (MF) of a neuron. The MF vector is of size k and each value of this vector is multiplied with the output weight and then summed up to get the computed output.

2.1. The Derivation of the Model of RBFN

We will derive an input-output relation underlying the architecture of RBFN in **Figure 1** in which prototype refers to the cluster centre. In this architecture there are two phases. The first phase is fuzzification and second phase is regression. For the fuzzification let us assume a cluster consisting of feature vectors of dimension k. Let i^{th} feature X_{ui} in this vector X_u be fuzzified using the i^{th} membership function μ_i and u stands for u^{th} input vector-output pair. Thus we have k fuzzy sets. Here we have as many neurons as the number of the input feature values. We don't require any equation for this phase. In the regression phase we employ Takagi-Sugeno-Kang fuzzy rule [8] on k-input fuzzy sets and one output as:

If X_{u1} is A_1 and X_{u2} is A_2 and $\cdots$ X_{uk} is A_k then

$$Y_u = b_0 + b_1 X_{u1} + b_2 X_{u2} + \cdots + b_k X_{uk} \tag{1}$$

where the fuzzy set $A_i = \{X_{ui}, P_{ui}\} \mid u = 1, \cdots, m; i = 1, 2, \cdots, k$. Now substituting the fuzzified inputs, *i.e.* $P_{ui} = \mu_i(X_{ui})$, for the inputs we get

$$Y_u = b_0 + b_1 P_{u1} + b_2 P_{u2} + \cdots + b_{uk} \tag{2}$$

This equation is valid if there is one class. We now extend this equation to the multi-class case. We feed the input vector of size k denoted by X_u and the neurons compute the membership function values P_{uj}. Let the number of classses be c. The regression equation that computes the outputs Y_l in multi-class is framed as:

$$Y_l = w_{0l} + w_{1l} P_{u1} + w_{2l} P_{u2} + \cdots + w_{kl} P_{uk};\ \mathbf{l}, = \quad \cdots, c \tag{3}$$

where we have replaced the weight vector $\{b_i\}$ by the weight matrix $\{w_{il}\}$ to account for multi-class. This is the governing equation for the architecture in **Figure 1**. The calculation of the output weights is deferred to Section III. The case when a class is represented by more than one cluster is now explained.

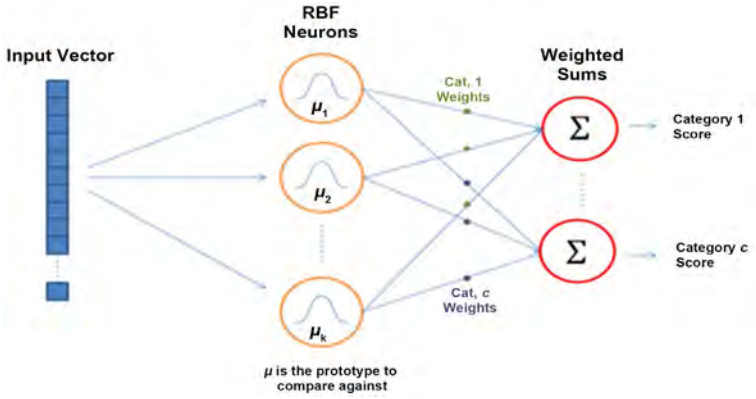


Figure 1. Architecture of RBFN.

2.2. Procedure for Learning of Weights

Consider first the problem of Iris flower recognition. In this we have 4 features. That is each feature vector is 4 dimensional. Assume that these are clustered into say, 3. It means that each cluster contains some number of feature vectors. According to the fuzzy set theory, we can form 4 fuzzy sets in each cluster corresponding to four features. Now each fuzzy set is defined by its attribute values and their membership function values. As we are using clustering, we can obtain mean values as well as scaling factors that are functions of variances involved in MFs of the fuzzy sets resulting from clustering. Our attempt is to focus on learning of weights.

Let us assume that we are feeding each feature vector of a cluster. Then we will have four neurons that convert four feature values of the feature vector into four membership function values. Then these membership values will be summed up. As we have assumed three clusters for one class (each flower type of Iris), this procedure is repeated on all feature vectors of the remaining two clusters. By this, we get three sums which will be multiplied with three weights (*i.e.* forming one weight vector) and the weighted sum is the computed output that represents one class.

The above procedure is repeated for the other three classes and the three weight vectors so obtained correspond to the remaining three flower types. There will also be three weighted sums called the computed outputs. In this paper, we are concerned with one cluster per class for simplicity.

3. Training of RBFN

The training process for RBFN consists of finding the three sets of parameters: the centrods of clusters, scaling parameters for each of the neurons of RBFN, and a set of the output weight vectors between the neurons and the output nodes of RBFN. The approaches for finding the centriods and their variances are discussed next.

3.1. Cluster Centrods

The possible approaches for the selection of clustercentrods are: Random

selection of centroids, Clustering based approach and Orthogonal Least Squares (OLS). We have selected the K-means clustering for the computation of centroids of clusters or cluster centres from the training setr. Note that this clustering method partitionsn observations into K number of clusters such that each observation having its value closest to the cluster centre belongs to that cluster.

3.2. Scaling Parameters

Camped with the centoid of each cluster, the variance is computed as the average distance between all points in the cluster and the centrod.

$$\sigma_i^2 = \frac{1}{m}\sum_{u=1}^{m}\left(X_{ui} - C_i\right)^2 \tag{4}$$

Here, C_i is the centroid of i^{th} cluster, m is the number of training samples belonging to this cluster, X_{ui} is the u^{th} training sample in the i^{th} cluster. Next we use σ_i^2 to compute the scaling parameters denoted by $\beta_i = \frac{1}{\sigma_i^2}$.

3.3. Output Weights

In the literature, there are two popular methods for the determination of the output weights: one learning method called gradient descent [11] and another computational method called pseudo inverse [12] [13]. As gradient descent learning has problems of slow convergence to local minima, we embark on a new learning algorithm called JAYA. Prior to using JAYA for learning the parameters of RBFN, we will discuss how the weights can be determined by Pseudo-inverse (PINV) method.

Consider an input vector which is generally a feature vector of some dimension n. When all the feature vectors are clustered, we will have C number of clusters (Note that c denotes the number of classes). In some datasets such as Iris dataset, we can easily separate out all the feature vectors belonging to each class of one flower type. Thus the feature vectors belonging to a class form a cluster. Out of these feature vectors some are selected for training and the rest for testing.

Let $\{X_u, Z_u\}; u = 1, \cdots, m$ be the set of feature vectors with each feature vector having the size of *n*, *i.e.* $X_u \in R^k$ with target, $Z_{ul} \in R^c$, and $P_{uj} = \mu_i\left(X_{uj}\right)$ be the membership function of the j^{th} basis radial function μ_j with the u^{th} feature vector. X_{uj} is the j^{th} component of the feature vector X_u and Z_{ul} is the l^{th} target output. Note that this formulation is meant for one cluster per one class. After the fuzzification of X_{uj} into P_{uj}, we can form a matrix P by taking $u = 1, \cdots, m$; $j = 1, 2, \cdots, k$. The matrix Q is written as

P_{11}	P_{12}	...	P_{1k}
P_{21}	P_{22}	...	P_{2k}
	..	...	..
P_{m1}	P_{m2}	...	P_{mk}

As we have $W_l = [w_{1l}, w_{2l}, \cdots, w_{kl}]$; $Z_l; l = 1, 2, \cdots, c$; $Z = [Z_1, Z_2, \cdots, Z_c]$. Let us denote $Q = [P_1, P_2, \cdots, P_m]$ with $P_u = [P_{u1}, P_{u2}, \cdots, P_{uk}]$ and $W = [W_1, W_2, \cdots, W_k]$. The objective function to be minimized is given by:

$$f(W, Q, Z) = \|W * Q - Z\|^2 \tag{5}$$

where $Y = W * Q$ with $Y = [Y_1, Y_2, \cdots, Y_c]$ as per Equation (3). The solution to the above equation lies in the assumption that be $Y = W * Q = Z$ which leads to $W = Q^+ Z$ where Q^+ denotes the pseudo inverse matrix of Q, defined as follows:

$$Q^+ = \lim_{\alpha \to 0} \left(Q^{\mathrm{T}} Q + \gamma I_n \right)^{-1} Q^{\mathrm{T}} \tag{6}$$

where I_k is the k-dimensional unity matrix and γ is a small positive constant. The pseudo inverse $Q^+ = \left(Q^{\mathrm{T}} Q \right)^{-1} Q^{\mathrm{T}}$ exists if $(Q^{\mathrm{T}} Q)$ is nonsingular. After calculating the weights at the output layer, all the parameters of RBFN with its 3-layered architecture in **Figure 1** can be determined.

4. Learning of the Output Weights by JAYA

We will now discuss JAYA algorithm to be used for learning the parameters of RBFN.

Description of the JAYA Algorithm

It is a simple and powerful learning method for solving the constrained and unconstrained optimization problems. As mentioned above JAYA algorithm is the offshoot of Teacher-Learner Based Optimization (TLBO) algorithm proposed in [14] [15]. This needs only the common controlling parameters like population size and number of iterations. The guidelines for fixing these parameters can be seen in [15]. Here we have fixed the population size as 10 and the number of iterations as 3000.

Let $f(W, Q, Z)$ be the objective function to be minimized. Let the best candidate be the one associated with the least value of the function (*i.e.* $f_{best}(W, Q, Z)$) and the worst candidate is the one with the highest value of the function (*i.e.* $f_{worst}(W, Q, Z)$) in all the candidate solutions. We choose B to stand for the weights W when the cluster centres and scale parameters are found separately. In case we use to learn all the parameters, B includes the cluster centres, scaling parameters and the output weights, *i.e.* $B = [C, \beta, W]$.

At any run of the algorithm, assume that there are 'j' design variables and 'k' candidate solutions and 'i' iterations. So to fit B into the JAYA algorithm, it is denoted by $B_{j,k,i}$ which is the value of the j^{th} variable of the k^{th} candidate during the i^{th} iteration. $B_{j,k,i}$ is updated to $B'_{j,k,i}$ during the iteration as,

$$B'_{j,k,i} = B_{j,k,i} + r_{1,j,i}\left(B_{j,best,i} - \left| B_{j,k,i} \right| \right) - r_{2,j,i}\left(B_{j,worst,i} - \left| B_{j,k,i} \right| \right) \tag{7}$$

where $B_{j,best,i}$ is the value of the j^{th} variable for the best candidate, $B_{j,worst,i}$ is the value of the j^{th} variable for the worst candidate at i^{th} iteration and $r_{1,j,i}$ and $r_{2,j,i}$ are the two random numbers in the range 0 to 1 for the j^{th} variable at the i^{th} iteration. The term $r_{1,j,i}\left(B_{j,best,i} - \left| B_{j,k,i} \right| \right)$ indicates the tendency to move closer

to the best solution whereas the term $r_{2,j,i}\left(B_{j,worst,i}-\left|B_{j,k,i}\right|\right)$ indicates the tendency to avoid the worst solution. $B'_{j,k,i}$ is accepted if its function value is better than that of $B_{j,k,i}$. All the accepted function values at the end of iteration are retained and these values become the input to the next iteration. The flowchart of JAYA algorithm is shown in Figure 2. Unlike TLBO algorithm that has two phases (*i.e.* teacher and learner), JAYA algorithm has only one phase and it is comparatively simpler to apply. Rao *et al.* [16] have used TLBO algorithm in the machining processes. A tea-category identification (TCI) system is developed in [17] and it uses a combination of JAYA algorithm and fractional Fourier entropy on three images captured by a CCD camera. In two studies involving heat transfer and pressure drop, *i.e.* thermal resistance and pumping power, two objective functions are used to ascertain the performance of the micro-channel heat sink. Multi-objective optimization aspects of plasma arc machining (PAM), electro-discharge machining (EDM), and micro electro-discharge machining (μ-EDM) processes are investigated in [18]. These processes are optimized while solving the multi-objective optimization problems of machining processes using MO-JAYA algorithm.

There are three learning parameters, viz., the cluster centers C_i, the scaling parameters (β_i) and the output weights W between the hidden and output layers. The learning of these parameters is depicted in Figure 3. The first parameter is found using K-means clustering algorithm.

We make use of JAYA algorithm for learning the second parameter, β_i. The weights are learned by optimizing the objective function using JAYA algorithm. The RBFN model so obtained can then be used for both classification and function approximation.

5. Design of Hanman Entropy Network

As the RBFN is not geared up to take care of the uncertainty in the input which may be an attribute or property value, we will make use of the Information set

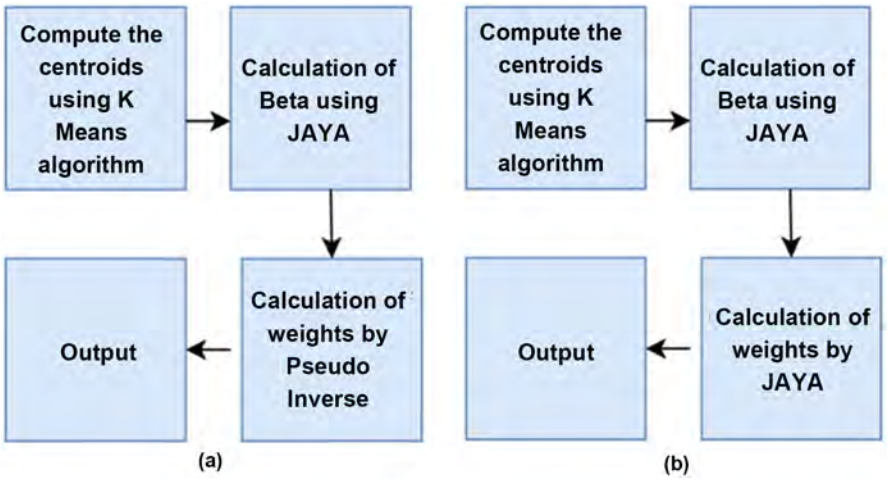


Figure 2. The learning processes in RBFN. (a) Weights using pseudo inverse; (b) Weights using JAYA Algorithm.

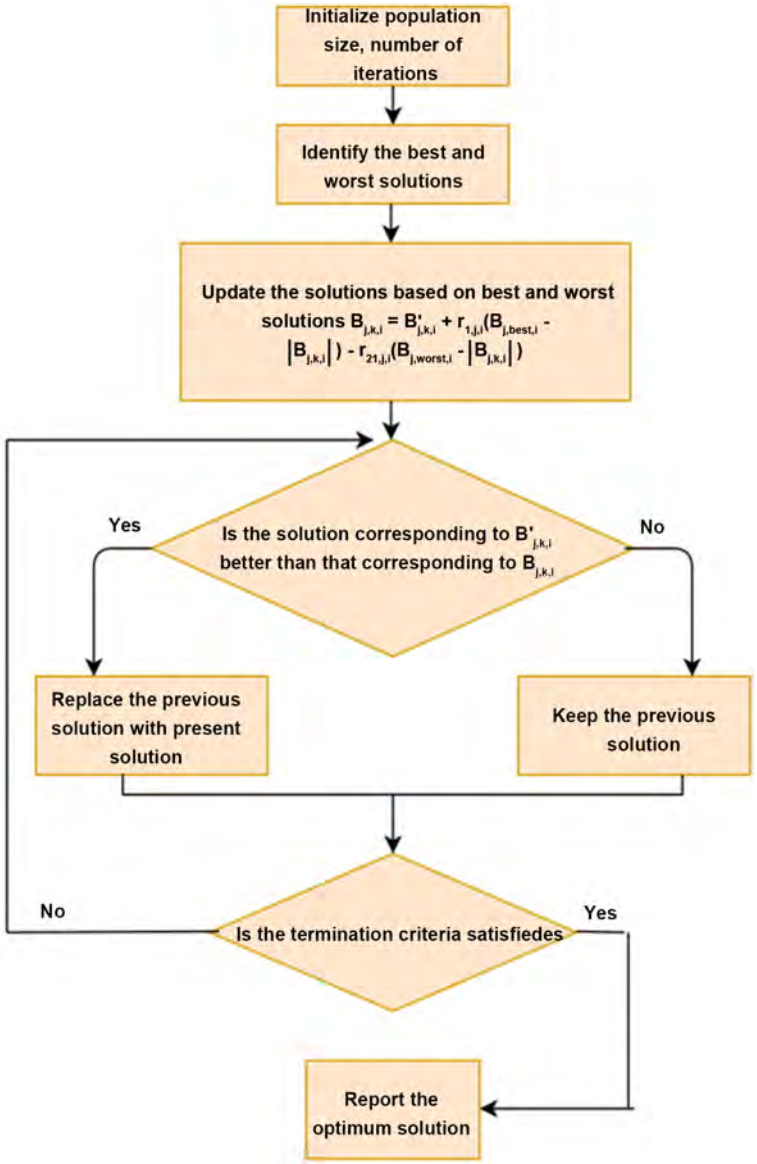


Figure 3. Flowchart of JAYA algorithm.

theory expounded by M. Hanmandlu and his co-workers [19] [20] [21]. In this theory each input variable is termed as the information source value. It centres around the concept of information set that expands the scope of a fuzzy set in which each element is a pair comprising a property (Information source) value and its degree of association with the set/concept called the Membership function value. In most of the applications involving fuzzy theory only the membership function is at the centres tage of operations. This limitation is sought to be removed by proposing the concept of information set. In real life contexts, we operate on information values. The information sources received by our senses are perceived by the mind as information values. That is the reason why we fail to perceive sound even when it strikes our ears. Like fuzzy variables, information values are also natural variables.

5.1. Definition of Information Set and Generation of Features

Consider a fuzzy set constructed from the feature values $\{X_{uj}\}$ termed as the Information source values and their membership function values which we take as the Gaussian function values $\{P_{uj}\}$. If the information source values do not fit the Gaussian function, we can choose any other mathematical function to describe their distribution. Thus each pair (X_{uj}, P_{uj}) consisting of information source value and its membership value is an element of a fuzzy set. P_{uj} gives the degree of association of X_{uj} and the sum of P_{uj} values doesn't provide the uncertainty associated with the fuzzy set. In the fuzzy domain, only P_{uj} is used in all applications of fuzzy logic thus ignoring X_{uj} altogether. This limitation is eliminated by applying the information theoretic entropy function called the Hanman-Anirban function [22] to the fuzzy set. This function combines the pair of values X_{uj} and P_{uj} into a product termed as the information value given by

$$H_{uj} = X_{uj} P_{uj} \tag{8}$$

The above relation owes its derivation to the non-normalized Hanman-Anirban entropy function expressed as

$$H = \sum_{j=1}^{n} X_{uj} \mathrm{e}^{-\left(aX_{uj}^3 + bX_{uj}^2 + cX_{uj} + d\right)} \tag{9}$$

where a, b, c and d are the real-valued parameters. In this equation normalization by n is not needed as the number of attributes is very small (less than 10 in the databases used) but needed if the value of H exceeds more than 1.

With the choice of parameters: $a = 0$, $b = \dfrac{1}{2\sigma^2}$, $c = -\dfrac{2\bar{X}}{2\sigma^2}$ and $d = \dfrac{\bar{X}^2}{2\sigma^2}$ where $\bar{X}$ is the mean value and σ^2 is the variance of the information source values X_{uj}, the exponential gain function is converted into the Gaussian function P_u. As a result, Equation (9) is modified to

$$H = \sum_{j=1}^{n} X_{uj} P_{uj} \tag{10}$$

The set of information values constitutes the information set denoted by $\mathcal{H} = \{H_{uj}\} = \{X_{uj} P_{uj}\}$ whereas the corresponding fuzzy set is simply $\{X_{uj}, P_{uj}\}$. Consider another entropy function called Mamta-Hanman entropy function [20] which is a generalized form of Hanman-Anirban entropy function, expressed as:

$$H_{MH} = \sum X_{uj}^{\alpha} \mathrm{e}^{-\left(cX_{uj}^{\gamma} + d\right)^{\rho}} \tag{11}$$

Substituting $c = \dfrac{1}{\sigma}$ and $d = -\dfrac{\bar{X}}{\sigma}$ in (11) modifies H_{MH} to the following:

$$H_{MH} = \sum_{j=1}^{n} X_{uj}^{\alpha} G_{uj} \tag{12}$$

where G_{uj} is the generalized Gaussian function given by

$$G_{uj} = \mathrm{e}^{-\left\{\frac{X_{uj}^{\gamma} - \bar{X}}{\sigma}\right\}^{\rho}}$$

This function takes different shapes as we vary the value of ρ from 1 to 5.

Assuming each information value as a unit of information, we can derive several modified information sets. For instance application of sigmoid function on $X_{uj}^{\alpha}G_{uj}$ leads to

$$S = \sum_{j=1}^{n} \frac{1}{1+\mathrm{e}^{-X_{uj}^{\alpha}G_{uj}}} \tag{13}$$

In Equations (9)-(13) the information values are the ones inside the summation sign. Thus a family of information forms can be deduced from both Hanman-Anirban and Mamta-Hanman entropy functions for dealing with different problems. For the derivation of different forms of H and H_{MH} the readers may refer to [19] and [21] respectively.

5.2. The Hanman Transform and its Link to Intuitionistic Set

This is a higher form of information set. To derive this transform, we have to consider the adaptive form of Hanman-Anirban entropy function in which the parameters in the exponential gain function are taken to be varaibles. Assuming $a = b = d = 0$ and $c = P_{uj}$ in (9) we obtain the Hanman Transform [21]:

$$H_T = \sum_{j=1}^{n} X_{uj}\mathrm{e}^{-X_{uj}P_{uj}} = \sum_{j=1}^{n} X_{uj}\mathrm{e}^{-H_{uj}} \tag{14}$$

Note that the exponentional gain function in (14) is a function of the information value. This transform acts as an evaulator of information values based on the information values obtained on them. The higher form of information set $\left\{X_{uj}\mathrm{e}^{-H_{uj}}\right\}$ is recursive because r.h.s of (14) can be rewritten as

$$H_{uj}\left(new\right) = X_{uj}\mathrm{e}^{-H_{uj}(old)} \tag{15}$$

An interesting result termed as Shannon transform emerges from Hanman transform by changing the substitution such that $d = -1$ instead of 0 in (9) and then simplifying the resulting exponential function as follows:

$$H_{Sh} = \sum_{j=1}^{n} X_{uj}\mathrm{e}^{1-H_{uj}} = \sum_{j=1}^{n} X_{uj}\log H_{uj} \tag{16}$$

Let us consider the adaptive Hanman-Anirban entropy function involving the membership functions alone. Then we have

$$H = \sum_{j=1}^{n} P_{uj}\mathrm{e}^{-\left(a(.)P_{uj}^{3}+b(.)P_{uj}^{2}+c(.)P_{uj}+d(.)\right)} \tag{17}$$

This gives the uncertainty in the membership function values. This is useful when a mathematical function describing the information source values is not appropriate thus leading to error in the fuzzy modeling. Now with a particular substitution of values for $a(.) = b(.) = d(.) = 0$ and $c(.) = X_{uj}$. Equation (14) takes the form

$$H_T = \sum_{j=1}^{n} P_{uj}\mathrm{e}^{-H_{uj}} \tag{18}$$

This equation paves the way for the recursive membership function value. The r.h.s. of (18) without summation can be written as:

$$P_{uj}(new) = P_{uj}(old)\mathrm{e}^{-H_{uj}(old)} \tag{19}$$

On the lines of derivation of (19), we can have another derivation from (11) as follows:

$$P_{uj}(new) = P_{uj}^{\alpha}(old)\mathrm{e}^{-H_{uj}(old)} \tag{20}$$

This is a useful relation becuase it can be used to make RBFN adaptive by changing the membership function. At this juncture we can make an interesting connection between the modified membership function in Equation (19) and hesitancy function in the Intuitionistic fuzzy set [23]. The hesitancey function is defined as follows:

$$h_{uj} = 1 - P_{uj} - P_{ucj} \tag{21}$$

where $P_{ucj} = 1 - P_{uj}$, the complemetary of P_{uj}. The hesitancy function reflects the uncertainty in the modeling of P_{uj} and P_{ucj}. As Equation (19) bestows the way to evaluate P_{uj}, we can use the new values of P_{uj} and P_{ucj} in determining the updated value of h_{uj} as follows:

$$h_{uj}(new) = 1 - P_{uj}(new) - P_{ucj}(new) \tag{22}$$

where $P_{ucj}(new) = P_{ucj}(old)\mathrm{e}^{-H_{ucj}(old)}$ and $H_{ucj} = X_{uj}P_{ucj}$. We can use this hesitancy function for the design of a new network in future.

5.3. Properties of Information Set

We will now present a few useful properties of Information set.

1) In the information set, the product of the complementary membership function value and the information source value gives the complementary information value.

2) Information values are natural variables just as signals received by biological neuron from visual cortex after modification by synapse.

3) The information values can be modified by applying various functions to provide effective features.

4) Higher form of information values like Hanman Transform provides a better representation of the information source values.

5) The fuzzy rules can be easily aggregated using the information set concept.

5.4. The Architecture and Model of Hanman Entropy Network

We will now discuss the architecture of HEN in **Figure 4**.

The architecture of HEN is the same as that of RBFN but for the function φ_i, which assumes the specified form of an entropy function of the input. In HEN each n-input vector needs to be categorized into any one of "c" classes. The i^{th} function denoted by φ_i converts all the values of the input vector into the entropy function values. This will be clear if we consider the Takagi-Sugeno-Kang fuzzy rule for multi-class case:

If X_{u1} is A_1 and X_{u2} is A_2 and $\cdots$ X_{un} is A_n then

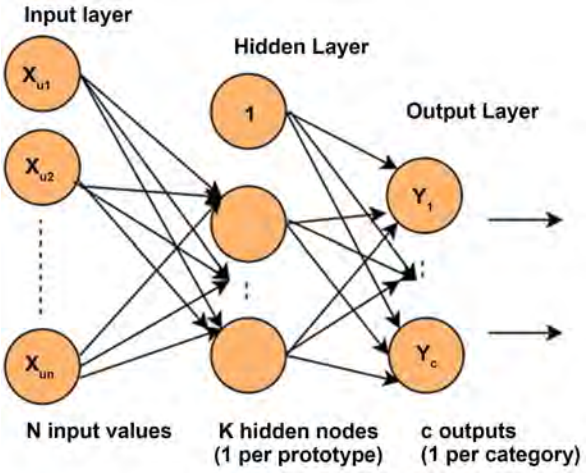


Figure 4. The Architecture of HEN.

$$Y_v = b_{v0} + b_{v1}X_{u1} + b_{v2}X_{u2} + \cdots + b_{vn}X_{un} \tag{23}$$

As mentioned above that at the fuzzification phase we replace b_{vj} with P_{vj} and set P_{v0} to 0 in (23) to get the neuron output as:

$$\varphi_v = P_{v1}X_{u1} + P_{v2}X_{u2} + \cdots + P_{vn}X_{un}; v = 1, \cdots \tag{24}$$

On substituting H_{uj} for the information values in (24) we get:

$$\varphi_v = H_{v1} + H_{v2} + \cdots + H_{vn}; v = 1, 2, \cdots, k \tag{25}$$

Thus the fuzzification phase in HEN is different from that of RBFN. In HEN the input feature vectors of size n are clustered into k clusters but in RBFN there is a single cluster for each feature of a feature vector. Each neuron in RBFN of Figure 1 has only one radial basis function whereas each neuron in HEN has k radial basis functions in Figure 4. So in Equation (25) k sums φ_1 to φ_k will be multiplied with the corresponding weights w_{1l} to w_{kl} to yield the l^{th} output Y_l in the regression phase as follows:

$$Y_l = w_{0l} + w_{1l} \cdot \varphi_1 + w_{2l} \cdot \varphi_2 + \cdots + w_{kl} \cdot \varphi_k; l = 1, \cdots, c \tag{26}$$

Here $\varphi_0 = 1$ and $\varphi = [1, \varphi_1, \varphi_2, \cdots, \varphi_k]$; $W_l = [w_{0l}, w_{1l}, \cdots, w_{kl}]$. This is the governing equation of Hanman Entropy Network in Figure 4. The objective function to be optimized by the JAYA + HEN combination is different. In view of this, the objective function becomes:

$$f(\varphi, W_l, Z_l) = \left\| W_l * \varphi - Z_l \right\|^2 \tag{27}$$

Instead of μ_j in RBFN, we will now use φ_j in HEN. In the general case φ_j can be taken as any relation linking the information source values to their membership function values. Thus, we can assume the following relations for this function:

$$\varphi_j = \sum_u \left\{ X_{uj}^{\alpha} G_{uj} \right\}_{\text{MH}} \quad \text{or} \quad \sum_u \left\{ \frac{1}{1 + e^{-X_{uj}^{\alpha} G_{uj}}} \right\}_{\text{S}} \quad \text{or} \quad \sum_u \left\{ X_{uj} e^{-H_{uj}} \right\}_{\text{T}} \quad \text{or}$$
$$\sum_u \left\{ X_{uj} \log H_{uj} \right\}_{\text{Sh}} \tag{28}$$

where subscripts MH, S, T and Sh indicate Mamta-Hanman, Sigmoid, Hanman transform and Shannon transform respectively. Note that RBFN network simply responds to the pattern in the input vector but the Hanman entropy network responds not only to the pattern but also to the uncertainty associated with it.

6. Results of Case Studies

The experimentation is conducted on four datasets: IRIS, Wine, and Waveform from UCI repository and Signature dataset [24] in two phases. In the first phase, we have entirely dealt with the performance analysis of RBFN and in the second phase only with that of Hanman Entropy Network (HEN). We have split up our computations into two cases. In Case-1 which is applicable only to RBFN, the learning/computation of the output weights is delinked from the computation of centriods and scaling parameters. Next, we have employed two learning methods such as Genetic algorithm (GA) and Gradient Descent (GD) and one computational method called Pseudo inverse (PINV) for the weights and K-means clustering for the centriods and scaling parameters.

There is another combination, JAYA + PINV + RBFN wherein JAYA is used for learning scaling parameters and PINV is used for computing the output weights. Of course, the centroids are found by K-means clustering.

The notations GA + RBFN, GD + RBFN, PINV + RBFN and JAYA +P INV refer to the learning of weights of RBFN by Genetic Algorithm, Gradient descent, Pseudo-Inverse, JAYA + PINV combination respectively. The results of classification accuracy with these four methods GA, GD, PINV and JAYA + PINV along with RBFN are given in **Table 1**. The last combination gives the best results. **Table 2** gives the comparison of JAYA + PINV + RBFN with JAYA + RBFN where the latter shows the best results.

A brief exposition on how to form fuzzy sets from which information sets are formed is the need of the hour. Assuming n feature types of an object, say, signature we form n fuzzy sets by collecting all the feature values of each feature type and fitting the radial basis function with the help of centroid (mean value of feature values) and scaling parameter (variance). Then conversion of this fuzzy set to information set is a simple matter.

The dataset-wise discussion of results follows the next.

IRIS dataset: This dataset consists of three classes with 50 samples for each class. There are four attributes for each sample. These are: sepal length, sepal

Table 1. A comparison of the Classification accuracy of RBFN using several learning methods on different datasets.

Dataset	GA + RBFN	GD + RBFN	PINV+RBFN	JAYA + PINV + RBFN
IRIS	96.3%	96.3%	96.3%	**96.7%**
WINE	86.88%	84.8%	82.6%	**92.13%**
WAVE FORM	87%	87.4%	87.5%	**85.8%**

Table 2. A comparison of the Classification accuracy of RBFN using JAYA for the learning of all parameters.

Dataset	JAYA + PINV + RBFN	JAYA + RBFN
IRIS	96.7%	**98%**
WINE	92.13%	**92.6%**
WAVEFORM	85.8%	**87.5%**

width, petal length and petal width all in cms. The classes are: Setosa, Versicolour and Virginica. It may be noted that GA, GD and PINV learning methods yield the classification accuracy of 96.3% when used for learning the weights of RBFN by classifying 144 instances out of 150. However, the combination symbolized by JAYA + PINV + RBFN gives the accuracy of 96.7% which uses 1) JAYA to learn the scaling parameters and 2) PINV to learn the weights of RBFN. This is the best result among all the results obtained on this dataset by the methods compared.

WINE dataset: The dataset consists of three classes with class 1, class 2, class 3 having 59, 71, 48 samples repectively. Each sample has 13 attribultes that include Alcohol, Malic acid, Ash, Alcalinity of ash, Magnesium, Total phenols, Flavanoids, Nonflavanoid phenols, Proanthocyanins, Color intensity, Hue, OD280/OD315 of diluted wines and Proline. The classification accuracies of 84.8% and 82.6% are achieved with GD and PINV respectively when these are used for the determination of the output weights of RBFN whereas GA + RBFN combination gives an accuracy of 86.88%. However, the best accuracy of 92.13% is achieved with JAYA-RBFN combination.

WAVEFORM dataset: The dataset consists of 5000 samples with each sample comprising 22 attributes. Each class is generated from a combination of 2 out of 3 "base" waves and each instance is generated by adding noise (mean 0, variance 1) to every attribute. RBFN classifies 4350 instances correctly out of 5000 instances with an accuracy of 87% with GA. The weights of RBFN computed using GD and PINV yield the best accuracies of 87.4% and 87.5% respectively but with JAYA + RBFN combimation the accuracy comes down to 85.8% in Table 2.

As can be seen from the results, the efficiency of the classification task increases when we use JAYA algorithm even for learning the weights of the network in comparison to learning the scaling parameters of the membership function. When the concept of information set is incorporated into our approach, the output is computed as Weights*Information values, *i.e.* $W_l * \varphi$. If the parameter vector, B also includes the centrods and the scaling parameters in addition to the weights then these parameters modify P_u indirectly. Then we will write $\varphi' = \varphi(C, \beta)$. Accordingly the objective function is modified as:

$$f(\varphi', W_l, Z_l) = \left\| W_l * \varphi' - Z_l \right\|^2 \tag{29}$$

For, RBFN the above with $P_i' = P_i(C, \beta)$ is written as

Table 3. A comparison of Verification accuracy on Signature dataset.

Dataset	JAYA + PINV + RBFN	JAYA + RBFN	HEN
SVC2004	90.7%	99.6%	99.8%

Table 4. Comparison of Classification accuracy.

Dataset	JAYA + RBFN	JAYA + HEN	MLP
IRIS	98%	**98.7%**	**95.3%**
WINE	92.6%	**93.2 %**	**81.4%**
WAVE FORM	87.5%	**88.6%**	**87.1%**

$$f\left(P_i', W_l, Z_l\right) = \left\| W_l * P_i' - Z_l \right\|^2 \tag{30}$$

Applying JAYA on (29) and (30) learns B.

SIGNATURE dataset: This dataset (SVC2004) in [24] has been used for a competition and it consists of 20 skilled forgeries and 20 genuine signatures of 40 users. Each signature in the dataset is represented as a sequence of points containing X and Y co-ordinates, time stamp and pen status (pen up or down) along with the additional information like azimuth, altitude and pressure. The text file contains a sequence of 7-dimensional measurements (feature types) for each signature. Our previous work on signature verification using Information set features on this dataset shows the effectiveness of these features [25]. We have used JAYA for learning both scaling parameters and the output weights of Hanman Entropy network (HEN) just as in JAYA + RBFN combination. The results of classification accuracy with JAYA + HEN are slightly better than those of JAYA + RBFN. But with JAYA + PINV + RBFN combination the results are very poor as shown in **Table 3**. The power of Hanman Entropy network can only be realized when the dataset is very large.

On conducting tests on three datasets as shown in **Table 4**, we find that the performance of JAYA + RBFN combination is somewhat inferior to that of JAYA + HEN combination on three datasets (Iris, Wine and Waveform) but the performance of Multi-layer perceptron (MLP) network is the worst. The use of high level information set features may help improve the performance of JAYA + HEN.

7. Conclusions

In this paper not only the performance of Radial Basis Function Network (RBFN) is improved by learning its parameters with a new evolutionary method called JAYA but also the design of Hanman Entropy network is given based on the Hanman-Anirban entropy function. Of all the combinations of RBFN with GA, GD, PNV, MLP and JAYA, JAYA + RBFN gives the best results. The proposed Hanman Entropy network (HEN) along with JAYA outperforms this combination on all the datasets considered in this paper.

As HEN is based on information set theory that caters to uncertainty representation; there is so much flexibility in the choice of information forms. This advantage is missing in RBFN where only the membership function values rule the roost. The only silver lining with RBFN is that we can use Type-2 fuzzy sets where the membership function values can be varied by changing the variance parameter of Gaussian membership function.

The present work opens up different directions to change the information at the hidden neurons of HEN.

Acknowledgements

The authors acknowledge the comments of anonymous reviewers, which helped improve the paper.

Conflicts of Interest

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

References

[1] Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) *Nature*, **323**, 533-536. https://doi.org/10.1038/323533a0

[2] Moody, J. and Darken, C.J. (1989) *Neural Computation*, **1**, 281-294. https://doi.org/10.1162/neco.1989.1.2.281

[3] Hecht-Nielsen, R. (1987) *Applied Optics*, **26**, 4979-4984. https://doi.org/10.1364/AO.26.004979

[4] Bullinaria, J.A. (2015) Radial Basis Function Networks: Algorithms, Neural Computation: Lecture 14. http://www.cs.bham.ac.uk/~jxb/INC/l14.pdf

[5] Poggio, T. and Girosi, F. (1990) *Proceedings of IEEE*, **78**, 1481-1497. https://doi.org/10.1109/5.58326

[6] Grover, J. and Hanmandlu, M. (2018) *Applied Intelligence*, **48**, 3394-3410. https://doi.org/10.1007/s10489-018-1154-x

[7] Rao, R.V. (2016) *International Journal of Industrial Engineering Computations*, **7**, 19-34. https://doi.org/10.5267/j.ijiec.2015.8.004

[8] Jang, J.S.R., Sun, C.T. and Mizutani, E. (2009) Neuro-Fuzzy and Soft-Computing: A Computational Approach to Learning and Machine Intelligence. PHI Learning Private Limited, New Delhi.

[9] Orr, M.J.L. (1995) *Neural Computation*, **7**, 606-623. https://doi.org/10.1162/neco.1995.7.3.606

[10] Xie, T.T., Yu, H. and Wilamowski, B. (2011) Comparison between Traditional Neural Networks and Radial Basis Function Networks. Industrial Electronics (ISIE). *IEEE International Symposium*, Gdansk, 27-30 June 2011, 1194-1199.

[11] Parappa, S.N. and Pratap Singh, M.P. (2013) *International Journal of Advancements in Research & Technology*, **2**, 112-125.

[12] Pierre, C. (2008) *Neural Information Processing—Letters and Reviews*, **8**, 25-29.

[13] Schwenker, F., Kestler, H.A. and Palm, G. (2001) *Neural Networks*, **14**, 439-458.

https://doi.org/10.1016/S0893-6080(01)00027-2

[14] Rao, R.V. and Patel, V. (2013) *Scientia Iranica*, **20**, 710-720. https://doi.org/10.1016/j.scient.2012.12.005

[15] Rao, R.V., Savsani, V.J. and Vakharia, D.P. (2011) *Computer-Aided Design*, **43**, 303-315. https://doi.org/10.1016/j.cad.2010.12.015

[16] Rao, R.V. and Kalyankar, V.D. (2012) *Journal of Materials and Manufacturing Processes*, **27**, 978-985. https://doi.org/10.1080/10426914.2011.602792

[17] Zhang, Y., Yang, X., Cattani, C., Rao, R.V. and Wang, S. (2016) *Entropy*, **18**, 77. https://doi.org/10.3390/e18030077

[18] Rao, R.V., Rai, D.P., Ramkumar, J. and Balic, J. (2016) *Advances in Production Engineering & Management*, **11**, 271-286. https://doi.org/10.14743/apem2016.4.226

[19] Hanmandlu, M. (2013) *Expert Systems with Applications*, **40**, 6478-6490. https://doi.org/10.1016/j.eswa.2013.05.020

[20] Hanmandlu, M. (2014) *Engineering Applications of Artificial Intelligence*, **36**, 269-286. https://doi.org/10.1016/j.engappai.2014.06.028

[21] Sayeed, F. and Hanmandlu, M. (2017) *Knowledge and Information Systems*, **52**, 485-507.

[22] Hanmandlu, M. and Das, A. (2011) *Defence Science Journal*, **61**, 415-430. https://doi.org/10.14429/dsj.61.1177

[23] Xu, Z. (2007) *IEEE Transactions on Fuzzy Systems*, **15**, 1179-1187. https://doi.org/10.1109/TFUZZ.2006.890678

[24] Chang, W. and Shin, J. (2008) DPW Approach for Random Forgery Problem in Online Handwritten Signature Verification. 4*th International Conference on Networked Computing and Advanced Information Management*, Gyeongju, 2-4 September 2008, 347-352. https://doi.org/10.1109/NCM.2008.118

[25] Hanmandlu, M., Sayeed, F. and Shantaram, V. (2015) Online Signature Verification Using the Entropy Function. 2015 *IEEE AIPR Workshop*, Washington DC, 13-15 October 2015, 1-7. https://doi.org/10.1109/AIPR.2015.7444522

Journal of Modern Physics, 2019, 10, 1522-1531
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Time Intervals of the Electron Transitions between the Energy States in the Hydrogen Atom Calculated in a Non-Probabilistic Way

Stanisław Olszewski

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

How to cite this paper: Olszewski, S. (2019) Time Intervals of the Electron Transitions between the Energy States in the Hydrogen Atom Calculated in a Non-Probabilistic Way. *Journal of Modern Physics*, **10**, 1522-1531.
https://doi.org/10.4236/jmp.2019.1013101

Received: October 15, 2019
Accepted: November 12, 2019
Published: November 15, 2019

Abstract

Definitions of the mechanical parameters entering the Bohr model of the hydrogen atom allowed us to calculate the time intervals connected with the electron transitions between the nearest-neighbouring energy levels in the atom. This is done in a strictly non-probabilistic way. The time results are compared with those derived earlier on the basis of the classical Joule-Lenz law for the energy emission adapted to the case of the electron transfer in the quantum systems. A similar formalism has been next applied to the harmonic oscillator and a particle moving in the one-dimensional potential box.

Keywords

Spectrum of the Hydrogen Atom, Time Intervals of the Electron Transitions

1. Introduction

Following the fundamental paper by Einstein [1] [2], a probabilistic approach to the intensity of the electron transitions in the atomic systems has been mainly applied: this has been done equally in the old as well as the modern quantum theory [2] [3] [4]. Physically this situation seems to be rather an encumbering one because the time intervals which are neglected are, in general, measurable parameters the size of which seems to be of interest for different processes.

A step towards a non-probabilistic theory of the electron transitions in the quantum systems has been done by the author [5] [6] [7]: it concerns the energy emission, therefore a transfer from a higher to a lower energy level, based on the classical Joule-Lenz law adapted to the quantum systems. When the transfer concerns solely the nearest-neighbouring quantum levels, say $n + 1$ and n, the

formula coupling the emitted energy ΔE to the emission time Δt becomes very simple, namely

$$\Delta E \Delta t = h; \tag{1}$$

here

$$\Delta E = E_{n+1} - E_n > 0. \tag{2}$$

Because of (1) and (2), the intensity of the energy emission from state $n+1$ to n is represented by the ratio

$$\frac{\Delta E}{\Delta t}. \tag{3}$$

The formula (3) could be next applied to the energy emission between the quantum states being also not nearest-neighbour ones. The numerical calculations were essentially limited to the ratios of the emission intensities for different kinds of the electron transitions in the hydrogen atom. The obtained results do not differ extensively from those calculated on the basis of the quantum-mechanical formalism; see [8] [9] [10].

Evidently (1) can be transformed into the formula representing Δt:

$$\Delta t = \frac{h}{\Delta E}. \tag{4}$$

In case of large n we have for the hydrogen atom

$$\Delta E = -\frac{me^4}{2\hbar^2}\left[\frac{1}{(n+1)^2} - \frac{1}{n^2}\right] \cong \frac{me^4}{2\hbar^2} \cdot \frac{2n+1}{n^4} \approx \frac{me^4}{\hbar^2 n^3}, \tag{5}$$

so in this case

$$\Delta t = \frac{h}{\Delta E} \cong \frac{h\hbar^2 n^3}{me^4} = \frac{2\pi\hbar^3 n^3}{me^4}. \tag{6}$$

The aim of the present paper is to demonstrate that results for Δt much similar to (6) can be easily obtained also by a direct application of the mechanical parameters entering the Bohr atomic model [11].

In **Appendix A** and **Appendix B**, respectively, the same formalism is applied to the harmonic linear oscillator and a particle enclosed in the one-dimensional potential box.

2. The Time Interval Δt Deduced from the Acceleration and Velocity Parameters Characteristic for the Bohr Atomic Model

The main idea by Bohr in constructing his model of the hydrogen atom (see e.g. [11] [12]) was that the centrifugal acceleration force of the electron due to its circulation about the proton nucleus is compensated by the electrostatic attractive force acting between the proton and the electron. Therefore the mechanical acceleration is given by the formula

$$a_n = \frac{v_n^2}{r_n} \tag{7}$$

where

$$v_n = \frac{e^2}{n\hbar} \tag{8}$$

is the velocity size along the circular electron orbit (see e.g. [12]) and

$$r_n = \frac{n^2\hbar^2}{me^2} \tag{9}$$

is the orbit radius [12]. Expressions (8) and (9) substituted into (7) give

$$a_n = \frac{e^4}{n^2\hbar^2}\frac{me^2}{n^2\hbar^2} = \frac{me^6}{n^4\hbar^4}. \tag{10}$$

In course of the electron transition from level $n+1$ to level n the acceleration is changed by the interval whose absolute value is

$$\begin{aligned}\Delta a_n &= \left|a_{n+1} - a_n\right| = \left|\frac{me^6}{\hbar^4}\left[\frac{1}{(n+1)^4} - \frac{1}{n^4}\right]\right| \\ &= \frac{me^6}{\hbar^4}\frac{(n+1)^4 - n^4}{(n+1)^4 n^4} \approx \frac{me^6}{\hbar^4}\frac{2n2n^2}{n^8} = \frac{4me^6}{\hbar^4 n^5},\end{aligned} \tag{11}$$

on condition in the last steps of (11) a large n is only considered. Simultaneously the absolute change of the electron velocity becomes

$$\Delta v_n = \left|\frac{e^2}{\hbar}\left(\frac{1}{n+1} - \frac{1}{n}\right)\right| = \frac{e^2}{\hbar}\frac{n+1-n}{(n+1)n} \approx \frac{e^2}{\hbar n^2}, \tag{12}$$

where the result obtained in the last step holds for large n. Since we expect the approximate validity of the formula

$$\frac{\Delta v_n}{\Delta t} = \Delta a_n, \tag{13}$$

we obtain from (11)-(13) the time interval

$$\Delta t = \frac{\Delta v_n}{\Delta a_n} = \frac{e^2}{n^2\hbar}\frac{\hbar^4 n^5}{4me^6} = \frac{n^3\hbar^3}{4me^4}. \tag{14}$$

The ratio of the result calculated in (6) to that obtained in (14) becomes equal to a constant number:

$$\frac{2\pi n^3\hbar^3}{me^4}\cdot\frac{4me^4}{n^3\hbar^3} = 8\pi. \tag{15}$$

3. The Time Interval Δt Deduced from the Electron Velocity and Orbit Length Parameters of the Bohr Model

A still more simple calculation of Δt than in Section 2 is connected with the case when the changes of the orbit length and electron velocity are taken into account. Here the change of velocity due to the electron transition between the levels $n+1$ and n is given by the formula

$$\frac{2\pi\Delta r_n}{\Delta t} = \Delta v_n \tag{16}$$

where

$$\Delta r_n = r_{n+1} - r_n = \left[(n+1)^2 - n^2\right]\frac{\hbar^2}{me^2} \cong 2n\frac{\hbar^2}{me^2} \tag{17}$$

and

$$\Delta v_n \cong \frac{e^2}{\hbar}\frac{1}{n^2} \tag{18}$$

is a repetition of the result obtained in (12). In effect from (16) we have

$$\Delta t = \frac{2\pi\Delta r_n}{\Delta v_n} = 4\pi n\frac{\hbar^3}{me^2}\frac{\hbar n^2}{e^2} = \frac{4\pi\hbar^3 n^3}{me^4}. \tag{19}$$

This result for Δt differs from that calculated in (6) solely by the factor of 2.

4. Time Interval of the Electron Transition and the Virial Theorem

Let us examine here the time interval Δt in reference to the virial theorem. Because of virial we have the following relation between the kinetic and potential electron energy in the hydrogen atom for any state n:

$$2E_n^{\text{kin}} + E_n^{\text{pot}} = 0. \tag{20}$$

In effect of (20) the electron energy becomes

$$E_n = E_n^{\text{kin}} + E_n^{\text{pot}} = -E_n^{\text{kin}} = -\frac{m}{2}v_n^2 \tag{21}$$

which implies

$$\left|\frac{\Delta E_n}{\Delta t}\right| \cong \frac{m}{2}2v_n\frac{\Delta v_n}{\Delta t} = mv_n\Delta a_n = m\frac{e^2}{n\hbar}\frac{4me^6}{\hbar^4 n^5} = \frac{4m^2e^8}{\hbar^5 n^6} \tag{22}$$

due to the formulae (8) and (11). By substituting the result in (5) for ΔE_n into (22) we obtain

$$\Delta t = \frac{\Delta E_n}{4m^2e^8}\hbar^5 n^6 = \frac{me^4}{\hbar^2 n^3}\frac{\hbar^5 n^6}{4m^2e^8} = \frac{\hbar^3 n^3}{4me^4} \tag{23}$$

which is the interval Δt identical to that given in (14).

5. Discussion

It looks that neither Einstein nor his followers were able to calculate a definite size of Δt, the time interval associated with the electron transitions in the atom, which, in the present considerations, is limited to the hydrogen case. In fact, the former authors were occupied mainly with the time changes of the transition probabilities than the time interval itself.

Our present idea concerning the time calculation was associated with an application of the classical definitions of the mechanical parameters (orbital position, electron velocity and acceleration) entering the Bohr model of the hydrogen atom. All these definitions make reference to the time interval Δt of the electron transition. But Δt could be calculated earlier on the basis of the clas-

sical Joule-Lenz law ([13] and [5] [6] [7]) adapted to the electron transitions between the nearest-neighbouring energy levels of the atom; see the end of Section 1.

In fact the essence of the formulae attained for Δt with the aid of the mechanical parameters in the Bohr atom and the expression for Δt obtained from the Joule-Lenz law is the same. The results of both approaches differ solely by the constant number represented in (15) or the factor of 2 derived at the end of Section 3.

Evidently the present semiclassical calculations do not take into account the degeneracy of the energy levels in the atom with respect to the electron angular momentum, or electron spin. In consequence the selection rules for transitions connected with the changes of the angular momentum could not be obtained; see [14].

As the end step we present the result of the numerical calculations of Δt in the hydrogen atom. By taking the formula (6) we obtain for large n:

$$\Delta t = \frac{2\pi\hbar^3}{me^4} n^3 \cong 2\pi \frac{\left(1.06\times 10^{-27}\right)^3}{9.1\times 10^{-28}\times\left(4.8\times 10^{-10}\right)^4} n^3 \text{ sec} \cong 1.5\times 10^{-16} n^3 \text{ sec}. \tag{24}$$

In consequence the intensity of a single energy transition—considered also for large n—is

$$\frac{\Delta E}{\Delta t} \cong \frac{3\times 10^5}{n^6} \frac{\text{erg}}{\text{sec}} \tag{25}$$

because from (5):

$$\Delta E = \frac{me^4}{\hbar^2 n^3} \cong \frac{5\times 10^{-11}}{n^3} \text{erg}. \tag{5a}$$

Both cases represented in (24) and (25) consider the quantum transition

$$n+1 \to n \tag{26}$$

for $n \gg 1$.

The derivations of the time interval done—classically—in the paper indicate that the size of Δt does not depend appreciably on the origin of Δt specifically whether we consider the process of the energy emission ($n+1 \to n$), or a reversed process of the absorption of energy ($n \to n+1$).

A separate problem concerns the situation when the electron transition, say that of the energy emission between the levels $n+\alpha$ and n, viz.

$$n+\alpha \to n \tag{27}$$

has

$$\alpha > 1. \tag{28}$$

In this case Equations (1) and (6) joining the intervals Δt and ΔE do not apply.

In order to overcome this difficulty the energy differences between the levels

$$\begin{aligned} &n+1 \text{ and } n, \\ &n+2 \text{ and } n+1, \\ &n+3 \text{ and } n+2, \\ &\quad\vdots \\ &n+\alpha \text{ and } n+\alpha-1 \end{aligned} \tag{29}$$

should be separately considered. For any pair of states entering (29) we can apply the formulae (1) and (6) which give us a set of the time intervals representing transitions between the succesive pairs of the energy levels indicated in (29).

A sum of these time intervals listed respectively as

$$\begin{aligned} &\Delta t_1, \\ &\Delta t_2, \\ &\Delta t_3, \\ &\quad\vdots \\ &\Delta t_\alpha \end{aligned} \tag{30}$$

provides us with the time interval between the levels $n+\alpha$ and n indicated in (27). Examples of such situations for the energy emission in the hydrogen atom are examined in [8] [9] [10].

Conflicts of Interest

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

References

[1] Einstein, A. (1917) *Physikalische Zeitschrift*, **18**, 121.

[2] Van der Waerden, B.L. (1967) Sources of Quantum Mechanics. Dover, New York.

[3] Slater, J.C. (1960) Quantum Theory of the Atomic Structure. McGraw-Hill, New York.

[4] Schiff, L.I. (1968) Quantum Mechanics. 3rd Edition, McGraw-Hill, New York.

[5] Olszewski, S. (2015) *Journal of Modern Physics*, **6**, 1277. https://doi.org/10.4236/jmp.2015.69133

Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 162.

[6] Olszewski, S. (2016) *Quantum Matter*, **5**, 664. https://doi.org/10.1166/qm.2016.1360

[7] Olszewski, S. (2016) *Reviews in Theoretical Science*, **4**, 1. https://doi.org/10.1166/rits.2016.1066

[8] Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 827. https://doi.org/10.4236/jmp.2016.78076

[9] Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 1004. https://doi.org/10.4236/jmp.2016.79091

Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 2314.

[10] Olszewski, S. (2017) *Journal of Computational and Theoretical Nanoscience*, **14**, 4086. https://doi.org/10.1166/jctn.2017.6791

[11] Bohr, N. (1922) The Theory of Spectra and the Atomic Constitution. Cambridge University Press, Cambridge.

[12] Sommerfeld, A. (1931) Atombau und Spektrallinien. Vol. 1, 5th Edition, Vieweg, Braunschweig.

[13] Lass, H. (1950) Vector and Tensor Analysis. McGraw-Hill, New York. https://doi.org/10.1119/1.1932684

[14] Condon, E.U. and Shortley, G.M. (1970) The Theory of Atomic Spectra. University Press, Cambridge.

[15] Sommerfeld, A. (1949) Mechanik, Akademische Verlagsgesellschaft, Leipzig.

[16] Eyring, H., Walter, J. and Kimball, G.E. (1957) Quantum Chemistry. Wiley, New York.

Appendix A. Transition Time between the Energy Levels $n+1$ and n of the Harmonic Oscillaror

First we note that the averages of the kinetic and potential parts of the energy W of a linear harmonic oscillator, viz.

$$W = \frac{m}{2}v^2 + \frac{k}{2}x^2 \tag{A1}$$

are equal; m is the mass, k, the oscillator constant [15]. This result can be obtained by substituting

$$x = A\cos\omega t \tag{A2}$$

where A is the amplitude and ω is the circular frequency of the oscillator. In effect

$$v = \frac{\mathrm{d}x}{\mathrm{d}t} = -A\omega\sin\omega t \tag{A3}$$

gives in average

$$\frac{m}{2}\overline{v^2} = \frac{m}{2}A^2\omega^2\overline{\sin^2\omega t} = \frac{m}{4}A^2\omega^2, \tag{A4}$$

whereas

$$\frac{k}{2}\overline{x^2} = \frac{k}{2}A^2\overline{\cos^2\omega t} = \frac{k}{4}A^2. \tag{A5}$$

The (A4) and (A5) are equal due to the relation

$$\omega = \left(\frac{k}{m}\right)^{1/2}. \tag{A6}$$

Evidently the amplitude $A = A_n$ is different for any quantum state of energy

$$W = W_n \cong n\hbar\omega \tag{A7}$$

which holds on condition n is a large integer number.

By considering the averages in (A4) and (A5) we have

$$\bar{E}_{\mathrm{kin}} = \frac{m}{2}\overline{v_n^2} = \frac{m}{4}A_n^2\frac{k}{m} = A_n^2\frac{k}{4} = \frac{1}{2}W_n \cong \frac{1}{2}n\hbar\omega. \tag{A8}$$

In effect

$$A_n \cong \frac{(2n\hbar\omega)^{1/2}}{k} \tag{A9}$$

and

$$\overline{v_n} \cong \left(\overline{v_n^2}\right)^{1/2} \approx \left(\frac{1}{2}\right)^{1/2} A_n\left(\frac{k}{m}\right)^{1/2} = \left(\frac{1}{2}\right)^{1/2} A_n\omega. \tag{A10}$$

The change of $\overline{v}$ due to the transition from state $n+1$ to n leads to:

$$\Delta\overline{v_n} \approx \left(\frac{1}{2}\right)^{1/2}\left(A_{n+1} - A_n\right)\omega. \tag{A11}$$

On the other side the change of the average velocity $\overline{v}$ can be considered as a direct effect of the change of the amplitude A in course of sometime interval

Δt:

$$\Delta\overline{v}_n \cong \frac{\Delta A_n}{\Delta t} = \left(A_{n+1} - A_n\right)\frac{1}{\Delta t}. \tag{A12}$$

A comparison of (A11) with (A12) leads to relation

$$\left(\frac{1}{2}\right)^{1/2}\left(A_{n+1} - A_n\right)\omega = \left(A_{n+1} - A_n\right)\frac{1}{\Delta t} \tag{A13}$$

which can be transformed into

$$\Delta t = \frac{2^{1/2}}{\omega}. \tag{A14}$$

On the other hand, according to the Joule-Lenz law, we have

$$\Delta t = \frac{h}{\Delta E} = \frac{h}{\hbar\left(n+1-n\right)\omega} = \frac{2\pi}{\omega}. \tag{A15}$$

A similarity between the Δt obtained from the particle mechanics in (A14) and Δt calculated from the Joule-Lenz law in (A15) becomes evident.

Appendix B. Transition Time between the Nearest-Neighbouring Energy Levels of a Particle Enclosed in a One-Dimensional Potential Box

The parameters characterizing a free particle of mass m in a one-dimensional box having the length L are given e.g. in [16]. According to the kinetic energy formula the particle in a quantum level n has the velocity

$$v_n = \left(\frac{2E_n}{m}\right)^{1/2} = \frac{1}{2}\frac{nh}{mL} \tag{B1}$$

because the particle kinetic energy is

$$E_n = \frac{n^2h^2}{8mL^2}. \tag{B2}$$

Evidently the velocity change between the levels $n+1$ and n is

$$\Delta v_n = \frac{1}{m^{1/2}}\left[\left(2E_{n+1}\right)^{1/2} - \left(2E_n\right)^{1/2}\right]. \tag{B3}$$

The energy change between the neighbouring levels is

$$\Delta E = E_{n+1} - E_n = \frac{\left(n+1\right)^2 - n^2}{8mL^2}h^2 \cong \frac{2nh^2}{8mL^2} = \frac{nh^2}{4mL^2}; \tag{B4}$$

in the last steps the condition $n \gg 1$ is considered.

The ΔE in (B4) can be coupled with the energy change entering (B3) by the relation

$$\frac{\Delta E}{E_{n+1}^{1/2} - E_n^{1/2}} = \frac{E_{n+1} - E_n}{E_{n+1}^{1/2} - E_n^{1/2}} = E_{n+1}^{1/2} + E_n^{1/2} \approx 2E_n^{1/2} \tag{B5}$$

where the last step holds for large n.

The acceleration change Δa_n of the particle associated with the levels change

from n+1 to n is

$$\Delta a_n = \frac{\Delta v_n}{\Delta t} \tag{B6}$$

where Δt is an unknown time interval. This interval can be obtained when the balance of energy change due to the change Δa_n is compared with the energy decrease (B4). This gives the relation

$$\Delta a_n m2L = \frac{\Delta v_n}{\Delta t} m2L = \Delta E = E_{n+1} - E_n \tag{B7}$$

in which we assumed that the energy loss is produced in course of a full free-particle oscillation in the box along the path length

$$L + L = 2L. \tag{B8}$$

By substituting in (B7) the necessary parameters we obtain the following relation for Δt entering (B6):

$$\begin{aligned} \Delta t &= \frac{2Lm\Delta v_n}{\Delta E} = \frac{2Lm}{\Delta E}\frac{2^{1/2}}{m^{1/2}}\left(E_{n+1}^{1/2} - E_n^{1/2}\right) = \frac{2^{3/2}Lm^{1/2}}{E_{n+1}^{1/2} + E_n^{1/2}} \\ &\approx \frac{2^{3/2}Lm^{1/2}}{2E_n^{1/2}} = 2^{1/2}\frac{L^2 m}{nh}8^{1/2} = \frac{4L^2 m}{nh}. \end{aligned} \tag{B9}$$

The end result of (B9) can be compared with Δt obtained from the Joule-Lenz law [5] [6]:

$$\Delta t = \frac{h}{\Delta E} = h\frac{4mL^2}{nh^2} = \frac{4L^2 m}{nh}. \tag{B10}$$

In fact the (B9) and (B10) are the identical expressions.

It should be noted that identity concerns also the velocity in (B1) and that obtained on the basis of the distance (B8):

$$v_n = \frac{2L}{\Delta t} = 2L\frac{nh}{4L^2 m} = \frac{nh}{2Lm}. \tag{B11}$$

Journal of Modern Physics, 2019, 10, 1532-1547
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Hypothesis of Primary Particles and the Creation of the Big Bang and Other Universes

Slobodan Spremo

Mathematical Grammar School, Belgrade, Serbia
Email: slobodan.spremo@gmail.com

How to cite this paper: Spremo, S. (2019) Hypothesis of Primary Particles and the Creation of the Big Bang and Other Universes. *Journal of Modern Physics*, **10**, 1532-1547.
https://doi.org/10.4236/jmp.2019.1013102

Received: October 14, 2019
Accepted: November 17, 2019
Published: November 20, 2019

Abstract

In this paper, we have presented a new approach to the dynamics of hypothetical primary particles, moving at speeds greater than the speed of light in a vacuum within their flat spacetime, which is why we understood the reason why they have not been detected so far. By introducing a new factor, we have linked the space-time coordinates of primary particles, within different inertial frames of reference. We have shown that transformations of coordinates for primary particles with respect to different inertial frames of reference, based on this factor, constitute the Lorentz transformations. Utilizing this factor, we have set the foundations of primary particle dynamics. The results obtained for the dynamic properties of these particles are in accordance with the fundamental laws of physics, and we expect them to be experimentally verifiable. Likewise, due to their dynamic properties, we have concluded that the Big Bang could have occurred during a mutual collision of the primary particles, with a sudden speed decrease of some of these particles to a speed slightly greater than the speed of light in a vacuum, which would release an enormous amount of energy. Created in such manner, our Universe would possess a limit on the maximum speed of energy-mass transfer, the speed of light in a vacuum, which we will show after introducing the dynamic properties of these particles. Similarly, we have concluded that the creation of other universes, possessing a different maximum speed of energy-mass transfer, occurred during the collision of these particles as well, only by means of deceleration of some of these particles to a speed slightly greater than the maximum speed of energy-mass transfer in that particular universe.

Keywords

Big Bang, Flat Spacetime, Lorentz Transformations

1. Introduction

According to the hypothesis, primary particles move at speeds u much greater

than the maximum speed of energy-mass transfer in different universes *k*, and because of this they are in their flat spacetime, wherein the aforementioned space is homogeneous and isotropic and time is homogeneous. Simultaneously, their lower border speed equals *k*, *i.e.* they can only move at speeds that surpass the values of it $u > k$. In our Universe, the maximum speed of energy-mass transfer is $k = c$. We have shown what the basic kinematic and dynamic properties of these hypothetical particles look like, based on the newly introduced ξ factor which links space-time coordinates between different inertial frames of reference. It is widely known that our Universe originated in the Big Bang, but that modern physics cannot explain events preceding Planck time. We expect further development of this hypothesis to clarify the Big Bang itself.

In the following section of this paper, we will discuss potential kinematic and dynamic properties of hypothetical primary particles with regard to our Universe, in which the maximum speed of energy-mass transfer is the same as the speed of light in a vacuum.

The facts known thus far are the following:

- Particles that possess the rest mass can approach the speed of light by increasing their momentum and energy. Let's assume they move with $v < c$.
- The maximum speed of energy-mass transfer is the same as the speed of light in a vacuum *c*.

Newly described:

- Primary particles, whose speeds we will denote by *u*, are able to move faster than light $c < u \le u_p$, $\left(u_p \gg c\right)$.

The postulates on which the primary particle hypothesis is based on, extend the first two postulates of the special theory of relativity with the third:

1) The Principle of Relativity: The laws of physics are invariant in all inertial frames of reference.

2) The Principle of Constancy of the speed of light: The speed of light in a vacuum is the same for all observers, regardless of the motion of the light source or observer.

3) The speeds of the primary particles *u*, may possess values $c < u \le u_p$, $\left(u_p \gg c\right)$, that are independent of the choice of inertial frame of reference from which observations are being made.

2. Time Dilation for Primary Particles

Einstein's relativity of the notion of simultaneous occurrence of some events [1] can also be extended to primary particles moving faster than light.

We will imagine how observers from the "mobile" *S'* and the "stationary" *S* frames of reference see the time between two events in the world of primary particles, through an example of the time required for the primary particle to pass a certain distance from the point of departure to the primary particle reflector (**Figure 1**). Let us assume that *S'* moves in the positive direction of the *x*-axis at the speed of light *c*, relative to system *S*, because we want to find the factor

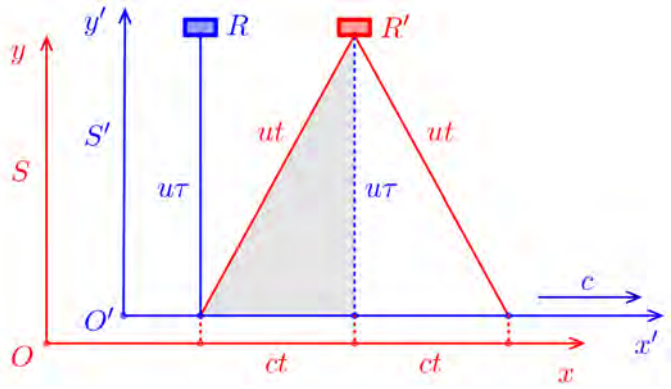


Figure 1. The observer *O* from the *S* perceives that the observer *O'* and the reflector are moving to the right at speed of light *c*, which represents the lower border velocity of the primary particles movement, hence the series of these events appear different to him. By the time primary particle reaches it, the reflector moves to the right by distance *ct*, hence the primary particle travels the distance *ut*, where *t* represents the time required for primary particle to relocate from the point of the viewer *O'* to the reflector at the position *R*. After time 2*t*, the primary particle, starting from the point *O'* and reflecting of the reflector, returns to *O'* again, but observed from the *S*, by passing the distance 2*ut*. Therefore, the observer *O* from the *S* concludes that the primary particle will reach the reflector only if it leaves the point of the observer at some angle in relation to the vertical plane. Note also that observer *O* must possess two synchronized clocks at the departure and arrival points of the primary particle, which are immobile in his frame of reference *S*, and compare their display with the display of a mobile clock located in the frame of reference *S'*. Thus, the observer *O* from the *S* measures the dilated time *t* in relation to the time τ measured by the observer *O'* from the *S'*.

through which the space-time connection of possible speeds of the primary particles *u* with *c*, would be implemented. Above the observer *O'* and the frame of reference *S'*, a reflector of primary particles *R'* is located at a certain distance, perpendicular to the path of the primary particle in that frame of reference. Observer *O'* measures time for which the primary particle proceeding perpendicular towards the reflector returned as 2τ, travels the distance $2u\tau$. Note that in order to measure this time, the observer *O'* requires only one clock which is permanently situated at the same location from which the primary particle starts and to which it returns.

From the shaded triangle, we can see the following:

$$u^2t^2 = c^2t^2 + u^2\tau^2, \quad t^2 = \frac{u^2}{u^2 - c^2}\tau^2, \quad t = \frac{1}{\sqrt{1-\frac{c^2}{u^2}}}\tau.$$

We can also introduce the ξ factor:

$$\xi = \frac{1}{\sqrt{1-\frac{c^2}{u^2}}}, \tag{1}$$

so the following applies $t = \xi\tau$

Similarly, if these events were to play out in the *S*, in the manner that the pri-

mary particle moved from and returned to the same place in this frame of reference, the observer *O'* from the *S'* would measure the same dilation of time, since in relation to him, the *S* moves at speed of light *c* only in the opposite direction, to the left. In accordance with the postulate on the independence of primary particle speed from choice of the inertial frame of reference from which it is being observed, for both observers it moves at the identical speed *u*. None of the observers *O* and *O'* have the means of determining whether or not they are moving, *i.e.* each of them is in a state of rest in his own frame of reference. Therefore, the names of two frames of reference "mobile" and "stationary" are written in inverted commas.

The shortest possible time between two events, proper time τ, is measured by the observer from whose perspective they are taking place at the same location in space. We therefore conclude that the time elapsed between the two events depends on how far they have unfolded within the two observed frame of references, *i.e.* that a connection between spatial and time intervals exists.

3. The Transformation of Space-Time Coordinates for Primary Particles

Since primary particles moving at speeds higher than the speed of light $u > c$, it is necessary to introduce new transformations of space-time coordinates between inertial frames of reference. They must apply to all speeds of these particles of *u* which is slightly higher than *c* to u_p. As we can see (**Figure 2**), the same event has different coordinates in two different inertial frames of reference.

$$x' = Cx + Dt, \quad t' = Kx + Lt, \tag{2}$$

where *C*, *D*, *K* and *L* represent constants to be determined. If we observe the movement of a primary particle along the *x*-axis, it will, from the perspective of the frame of reference *S'* at time moments t_1' and t_2', possess the spatial coordinates x_1' and x_2', and its shift will be

$$\Delta x' = x_2' - x_1' = C\left(x_2 - x_1\right) + D\left(t_2 - t_1\right). \tag{3}$$

The time interval in the *S'*, during which the movement occurred is

$$\Delta t' = t_2' - t_1' = K\left(x_2 - x_1\right) + L\left(t_2 - t_1\right). \tag{4}$$

Based on this, the velocity of the observed primary particle moving along the *x*-axis equals:

$$u' = \frac{\Delta x'}{\Delta t'} = \frac{C\Delta x + D\Delta t}{K\Delta x + L\Delta t}. \tag{5}$$

Relative to the frame of reference *S* the velocity of that same point is $u = \frac{\Delta x}{\Delta t}$, and the relation of these two velocities is given via expression

$$u' = \frac{Cu + D}{Ku + L}. \tag{6}$$

In order to determine the constants *C*, *D*, *K* and *L* in this expression, we will consider several special cases of motion.

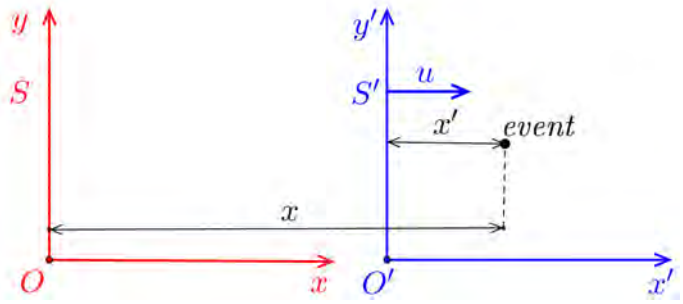


Figure 2. Let us assume the frame of reference *S* to be "stationary" and the frame of reference *S'* to be "mobile", moving in the positive direction of *x*-axis at a constant velocity of *u*. Assume that both reference systems coincided at the moment of time $t = t' = 0$. The same event would be described by the observer from *S* in space-time coordinates (x, y, z, t), while the observer from *S'* would assign them the coordinates (x', y', z', t'). It is necessary now to find new coordinate transformations that would also apply at the higher relative velocities of frames of reference, than the speed of light. Along the *y* and *z* axes, no movement *S* and *S'* occurs, and as in the cases of Galilean and Lorentz transformations, $y' = y$, $z' = z$, applies. In order to determine the functional connection between the remaining two coordinates, one of space and one of time, we have to acknowledge that it should preserve the properties of space and time. The basic properties of the space are homogeneity and isotropy, and the basic property of time is homogeneity, which is directly related with the conservation laws in mechanics. Therefore, the coordinate transformation law we are searching for has to be linear, hence the interconnection of the coordinates is linear.

We will first observe the border case. The primary particle is idle relative to *S'*, *i.e.* is located in its own frame of reference, at its lowest possible velocity in relation to *S*. In this case, $u' = 0$ while *u* has slightly greater value than *c*, *i.e.* $u \to c$. Therefore, observed from a "stationary" frame of reference, the primary particle moves at the same velocity as the "mobile" frame of reference. Thus, expression (6) becomes

$$0 = \frac{Cc + D}{Kc + L},$$

and the following applies

$$D = -Cc. \tag{7}$$

Vice versa, due to relativity of motion, when the primary particle is idle relative to *S i.e.* $u = 0$, then its velocity in relation to *S'*, $u' \to c$. If we substitute this in (6) using (7) we get $-c = -\frac{Cc}{L}$, *i.e.*

$$C = L. \tag{8}$$

We will now use the third postulate according to which the possible primary particle velocities appear identical, observed from all inertial frames of reference

$$u = u', \tag{9}$$

and if we substitute this in (6) using (7) and (8) we get $u = \frac{Cu - Cc}{Ku + C}$, from

which the following applies

$$K = -C\frac{c}{u^2}. \tag{10}$$

Equations (2), after replacing the values D, K and L, become

$$x' = C(x - ct), \quad t' = C\left(t - \frac{c}{u^2}x\right), \tag{11}$$

in which only the constant C remains unspecified. According to the principle of relativity, a complete equality of the observed frames of reference exists. Which means that for an "stationary" frame of reference we can take S' and consider that the frame of reference S moves at velocity "$-u$" in relation to it. Based on that, equations that connect x and t with x' and t' read

$$x = C(x' + ct'), \quad t = C\left(t' - \frac{c}{u^2}x'\right). \tag{12}$$

When we replace (12) in (11) we get

$$x' = C^2\left(x' + ct' - ct' - \frac{c}{u^2}x'\right), \tag{13}$$

from which we can see that $C = \dfrac{1}{\sqrt{1-\dfrac{c^2}{u^2}}}$, *i.e.* C have the same values as the previously introduced ξ factor (1). With constants determined in such a manner, we obtain the transformations in the form

$$x' = \frac{x - ct}{\sqrt{1-\dfrac{c^2}{u^2}}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t - \dfrac{c}{u^2}x}{\sqrt{1-\dfrac{c^2}{u^2}}}, \tag{14}$$

i.e.

$$x = \frac{x' + ct'}{\sqrt{1-\dfrac{c^2}{u^2}}}, \quad y = y', \quad z = z', \quad t = \frac{t' + \dfrac{c}{u^2}x'}{\sqrt{1-\dfrac{c^2}{u^2}}}. \tag{15}$$

These transformations connect the space-time coordinates (x, y, z, t) and (x', y', z', t') of the same event observed from two inertial frames of reference S and S' in the case of their motion being relative along x-axis at a constant velocity of $u > c$.

4. Four-Dimensional Formulation of Coordinate Transformations for Primary Particles

Similar to Minkowski space [2], we shall introduce a real four-dimensional space in which we will present transformations of coordinates of the primary particles as transformations of coordinates of that very space. Points in that space are position vectors

$$\boldsymbol{x} = x^{\sigma}\boldsymbol{e}_{\sigma} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ut \\ x \\ y \\ z \end{pmatrix}, \tag{16}$$

where x^{σ} represent contravariant components of the vector $\boldsymbol{x}$ at basis

$$\boldsymbol{e}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

The metric of this space is identical to the metric of Minkowski space

$$g_{\sigma\varsigma} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{17}$$

This metric is used to determine the length of a vector. A square of length of the position four-vector x equals

$$\begin{aligned} x^2 = x^{\mathrm{T}} g x &= \begin{pmatrix} ut & x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} ut \\ x \\ y \\ z \end{pmatrix} \\ &= \left(x^0\right)^2 - \left(x^1\right)^2 - \left(x^2\right)^2 - \left(x^3\right)^2 = u^2t^2 - \boldsymbol{x}^2. \end{aligned} \tag{18}$$

Within spacetime of the primary particles, we associate each contravariant vector to covariant components using the metric

$$A_{\sigma} = g_{\sigma\varsigma} A^{\varsigma}. \tag{19}$$

Below, we will use Einstein summation convention, which implies summation when an index is repeated twice in a single term, once as upper index and once as lower index, without writing the sum sign. Therefore, the previous form is summed by the index ς, and the following applies

$$\begin{aligned} A_0 &= g_{0\varsigma} A^{\varsigma} = g_{00}A^0 + g_{01}A^1 + g_{02}A^2 + g_{03}A^3 = A^0, \\ A_1 &= g_{1\varsigma} A^{\varsigma} = g_{10}A^1 + g_{11}A^1 + g_{12}A^2 + g_{13}A^3 = -A^1. \end{aligned} \tag{20}$$

In a similar way, we get $A_2 = -A^2$ and $A_3 = -A^3$. We see that lowering the time index does not change the sign, while lowering the spatial index $i = 1, 2, 3$ does change the sign, so we can write

$$A_0 = A^0, \quad A_i = -A^i. \tag{21}$$

The covariant components of the position vector are

$$\begin{aligned} x_0 &= x^0 = ut, \\ x_1 &= -x^1 = -x, \\ x_2 &= -x^2 = -y, \\ x_3 &= -x^3 = -z. \end{aligned} \tag{22}$$

An inverse metric tensor is defined by

$$gg^{-1} = g^{-1}g = I, \tag{23}$$

where I represents an identity matrix. Numerically, we easily get $g^{-1} = g$. However, tensors g^{-1} and g have different index layout. The inverse metric allows us to transform the lower indices into the upper ones *i.e.* $A^{\sigma} = g^{\sigma\varsigma}A_{\varsigma}$.

The square of the length of vector x is calculated via

$$x^2 = g_{\sigma\varsigma}x^{\sigma}x^{\varsigma} = \left(x^0\right)^2 - \boldsymbol{x}^2, \tag{24}$$

and for differential of the square length between the points x and $x + \mathrm{d}x$ we get

$$\mathrm{d}s^2 = g_{\sigma\varsigma}x^{\sigma}x^{\varsigma} = u^2\left(\mathrm{d}t\right)^2 - \left(\mathrm{d}\boldsymbol{r}\right)^2. \tag{25}$$

Boost along x-axis for primary particles, which is similar to the Lorentz boost, we will record in the form of

$$\begin{aligned} x'^0 &= \xi x^0 - \nu\xi x^1 \\ x'^1 &= -\nu\xi x^0 + \xi x^1 \\ x'^2 &= x^2 \\ x'^3 &= x^3 \end{aligned}$$

where the following applies:

$$\nu = \frac{c}{u}, \quad \xi = \frac{1}{\sqrt{1 - \dfrac{c^2}{u^2}}}. \tag{26}$$

In matrix form, the previous equations are

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \xi & -\nu\xi & 0 & 0 \\ -\nu\xi & \xi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{27}$$

where 4×4 matrix represents the transformation matrix $\Sigma_{\varsigma}^{\sigma}$, in which the index σ represents the row index and index ς is the column index. The matrix Equation (27) in component notation is

$$x'^{\sigma} = \Sigma_{\varsigma}^{\sigma}x^{\varsigma}. \tag{28}$$

Lorentz transformations are those linear transformations of the coordinates $x' = \Lambda x$, where Λ represents the real matrix 4×4, which does not alter the square of length of the four-vector, *i.e.* the following applies to them $x'^2 = x^2$, which gives

$$\Lambda^{\mathrm{T}} g \Lambda = g. \tag{29}$$

Therefore, every real 4×4 matrix that satisfies the condition (29) is a Lorenz transformation. Hence, we see that our boost matrix along x-axis is a Lorentz transformation, and it can be easily shown that the boost matrices along y-axis and z-axis are also Lorentz transformations, as well as the three matrices of rotation of the coordinate system.

5. Time Dilation as a Result of a Transformation of Coordinates for Primary Particles

We have already derived the time dilation formula for primary particles based on postulates that apply to primary particles. We will now demonstrate how it is obtained by applying the transformations of the coordinates of the primary particles. Imagine the clock being located at the coordinate start of the inertial frame of reference *S'*. It will show time *S'* in *t'* and its coordinates in *S* would be $x' = y' = z' = 0$. If we replace those values of time coordinates with expressions for transformation of the space-time coordinates of primary particles (15), we get the coordinates of the clock from "stationary" frame of reference *S* in relation to which *S'* moves at velocity *u*: $x = \frac{ct'}{\sqrt{1-\frac{c^2}{u^2}}}$, $y = 0$, $z = 0$, since, observed from the *S*, clock moves along *x*-axis. Thus, the remaining time coordinate

$$t = \frac{t'}{\sqrt{1-\frac{c^2}{u^2}}}.$$

As we have shown earlier, the time shown by the clock in frame of reference *S'*, in relation to which it remains immobile, equals less than the time measured in frame of reference *S*, that is, in this case, time *t'* represents own time.

6. The Velocity-Addition Formula for Primary Particles

We will continue to observe that *S* as a "stationary" frame of reference in relation to which *S'* moves at velocity $\boldsymbol{u} = (u_x, u_y, u_z)$. Primary particle with the velocity of $\boldsymbol{w} = (w_x, w_y, w_z)$ in relation to *S*, possess the velocity $\boldsymbol{w}' = (w_x', w_y', w_z')$ in relation to *S'*, while its *x* component

$$w_x' = \frac{\mathrm{d}x'}{\mathrm{d}t'}.$$

According to the Equation (14), the differential $\mathrm{d}x'$ and the differential $\mathrm{d}t'$ are

$$\mathrm{d}x' = \frac{\mathrm{d}x - c\mathrm{d}t}{\sqrt{1-\frac{c^2}{u^2}}}, \quad \mathrm{d}t' = \frac{\mathrm{d}t - \frac{c}{u^2}\mathrm{d}x}{\sqrt{1-\frac{c^2}{u^2}}}, \tag{30}$$

while the velocity represents their ratio

$$w_x' = \frac{\frac{\mathrm{d}x}{\mathrm{d}t} - c}{1 - \frac{c}{u^2}\frac{\mathrm{d}x}{\mathrm{d}t}}.$$

Since $w_x = \frac{\mathrm{d}x}{\mathrm{d}t}$, *x* component of velocity relative to *S*, we have

$$w_x' = \frac{w_x - c}{1 - \frac{w_x c}{u^2}}. \tag{31}$$

Correspondingly, starting from $w_y' = \frac{\mathrm{d}y'}{\mathrm{d}t'}$ and $w_z' = \frac{\mathrm{d}z'}{\mathrm{d}t'}$, for the remaining two components of the velocity we get

$$w_y' = \frac{w_y\sqrt{1-\frac{c^2}{u^2}}}{1-\frac{w_x c}{u^2}}, \quad w_z' = \frac{w_y\sqrt{1-\frac{c^2}{u^2}}}{1-\frac{w_x c}{u^2}}. \tag{32}$$

In the border case $w_x = u$ formula (31) becomes

$$w_x' = \frac{u-c}{1-\frac{uc}{u^2}} = \frac{u\left(1-\frac{c}{u}\right)}{1-\frac{c}{u}} = u.$$

From this we can conclude that if a primary particle moves at a speed of u relative to an observer from the frame of reference S, then, regardless of the relative speed of the system, the primary particle possesses the same speed in relation to the observer from the frame of reference S' as well. We introduced this within the third postulate.

Formulas (31) and (32) allow us to determine the components of velocity in relation to S' for known velocity components in relation to S. Vice versa, the following applies

$$w_x = \frac{w_x' + c}{1+\frac{w_x' c}{u^2}}, \tag{33}$$

$$w_y = \frac{w_y'\sqrt{1-\frac{c^2}{u^2}}}{1+\frac{w_x' c}{u^2}}, \quad w_z = \frac{w_z'\sqrt{1-\frac{c^2}{u^2}}}{1+\frac{w_x' c}{u^2}}. \tag{34}$$

7. Acceleration with Regard to the Primary Particle Hypothesis

Descartes' components of acceleration of the primary particle in inertial frame of reference S are defined as

$$a_x = \frac{\mathrm{d}w_x}{\mathrm{d}t}, \quad a_y = \frac{\mathrm{d}w_y}{\mathrm{d}t}, \quad a_z = \frac{\mathrm{d}w_z}{\mathrm{d}t}, \tag{35}$$

and in S' are

$$a_x' = \frac{\mathrm{d}w_x'}{\mathrm{d}t'}, \quad a_y' = \frac{\mathrm{d}w_y'}{\mathrm{d}t'}, \quad a_z' = \frac{\mathrm{d}w_z'}{\mathrm{d}t'}. \tag{36}$$

From (31) we obtain

$$\mathrm{d}w_x' = \frac{\mathrm{d}w_x}{\left(1-\frac{w_x c}{u^2}\right)^2}\left(1-\frac{c^2}{u^2}\right). \tag{37}$$

The rate of change of the x velocity component in S' is given via relation (30) we have

$$a'_x = \frac{\mathrm{d}w'_x}{\mathrm{d}t'} = \frac{\left(1-\frac{c^2}{u^2}\right)^{\frac{3}{2}}}{\left(1-\frac{w_x c}{u^2}\right)^3} a_x. \tag{38}$$

For the remaining two acceleration components using (32), an analogous method produces

$$a'_y = \frac{\mathrm{d}w'_y}{\mathrm{d}t'} = \frac{1-\frac{c^2}{u^2}}{\left(1-\frac{w_x c}{u^2}\right)^2} a_y + \frac{w_y c\left(1-\frac{c^2}{u^2}\right)}{u^2\left(1-\frac{w_x c}{u^2}\right)^3} a_x,$$

$$a'_z = \frac{\mathrm{d}w'_z}{\mathrm{d}t'} = \frac{1-\frac{c^2}{u^2}}{\left(1-\frac{w_x c}{u^2}\right)^2} a_z + \frac{w_z c\left(1-\frac{c^2}{u^2}\right)}{u^2\left(1-\frac{w_x c}{u^2}\right)^3} a_x. \tag{39}$$

Based on these relations, which enable transformation of the acceleration components between the two inertial frames of reference, we can conclude that acceleration is not invariant, similar to that in the special theory of relativity. However, all inertial observers will agree that acceleration exists. Namely, if a primary particle observed from an inertial frame of reference is moving rapidly, then it is moving rapidly in all other inertial frames of reference as well. Similarly, if the acceleration in one of those inertial frames of reference equals zero, then it equals zero in all other inertial frames of reference.

8. Fundamentals of Primary Particle Dynamics

While moving at maximum speeds $u_p \gg c$, energy E_p of these particles would match the fundamental rest energy in relativistic physics

$$E_p = m_p c^2, \tag{40}$$

where m_p represents the mass of the primary particle.

Similar to the total energy of a relativistic particle noted via Lorentz γ factor, we can use the ξ factor for notation of the total energy of a primary particle E_t as

$$E_t = \xi m_p c^2 = \frac{m_p c^2}{\sqrt{1-\frac{c^2}{u^2}}}. \tag{41}$$

From (41) we can observe that the total energy of the primary particles behaves depending on their speed as in **Figure 3**. They possess the lowest energy $E_t = m_p c^2$. while moving at maximum speed $u_p \gg c$, and the highest when $u \to c$.

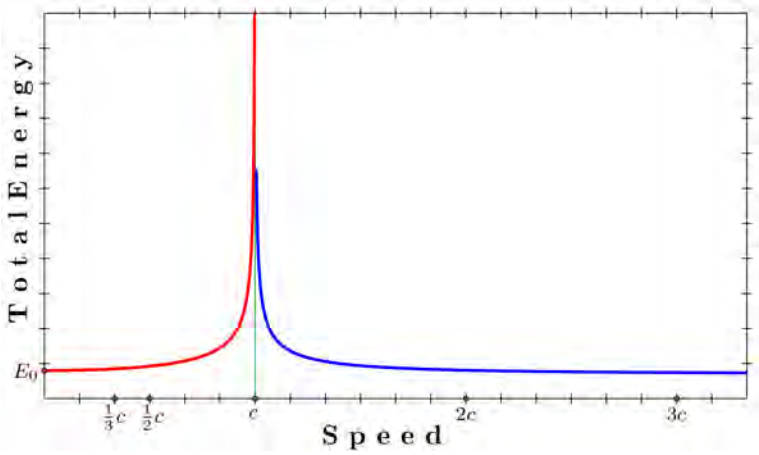


Figure 3. The dependence of total relativistic energy of the particle on speed v (red curve), the speed of light in a vacuum c (green line) and dependence total energy of the primary particle on speed u; according to the primary particles hypothesis (blue curve).

We will introduce homeokinetic energy *H*.

Conversely with kinetic energy increment that accompanies increase of the speed of particles possessing rest mass in classical and relativistic physics, homeokinetic energy of primary particles increases as they decelerate during mutual collisions, consequently increasing their total energy E_t.

$$E_t = E_p + H. \tag{42}$$

It is obvious that these particles would not interact with matter via four known interactions due to the speed of their motion $u > c$, so primary particles therefore exist in their flat spacetime, while the non-interaction of these particles with matter is the reason why they have not been detected so far. So, in order for these particles to mutually interact, a need for a new kind of interaction expanding faster than the speed of light arises.

For homeokinetic energy, from (40), (41) and (42) we get

$$H = E_t - E_p = m_p c^2 \left(\frac{1}{\sqrt{1-\frac{c^2}{u^2}}} - 1 \right) = m_p c^2 \left(\xi - 1 \right). \tag{43}$$

From the energy-momentum relation $E^2 - E_0^2 = p^2 c^2$, which is fundamental, similarly, for primary particles we have

$$E_t^2 - E_p^2 = p_p^2 c^2, \tag{44}$$

where p_p is the momentum of a primary particle. Here, considering (40) and (41), we see

$$m_p^2 c^4 \left(\frac{1}{\frac{u^2-c^2}{u^2}} - 1 \right) = \frac{m_p^2 c^4 - m_p^2 c^4 \frac{u^2-c^2}{u^2}}{\frac{u^2-c^2}{u^2}} = \frac{m_p^2 \frac{c^6}{u^2}}{\frac{u^2-c^2}{u^2}} = p_p^2 c^2,$$

from which we obtain the formula for the magnitude of the momentum a primary particle, differently defined compared to classical and relativistic physics (Figure 4)

$$p_p = \frac{m_p c^2}{u\sqrt{1-\frac{c^2}{u^2}}}, \tag{45}$$

so that:

$$p_p = \frac{E_t}{u}. \tag{46}$$

Also, the formula for the magnitude of the momentum a primary particle (46) can be obtained in a different, more intuitive way. In relativistic physics, the magnitude of the momentum a photon is $p = \frac{E}{c}$, where the speed of light in a vacuum c is independent of the choice of inertial frame of reference from which observations are being made. Similarly, the speeds of the primary particles u are also independent of the choice of inertial frame of reference from which observations are being made, hence the formula for the magnitude of the momentum a primary particle is $p_p = \frac{E_t}{u}$.

For $u \gg c$ the following applies: $p_p \to 0$.

We will further show that the basic dynamic laws of physics apply to this form of total energy of primary particles.

We will now determine the work W performed by the force F, directed along the x-axis, during deceleration of the primary particle relocating from a certain location determined by the parameters (x_1, t_1, u_p), to the second location determined by the parameters (x_2, t_2, u). u_p is the maximum speed of the primary particle ($u_p \gg c$). Using the definition of work and Newton's second law, which applies in the same form to primary particles as well, in accordance with the first postulate

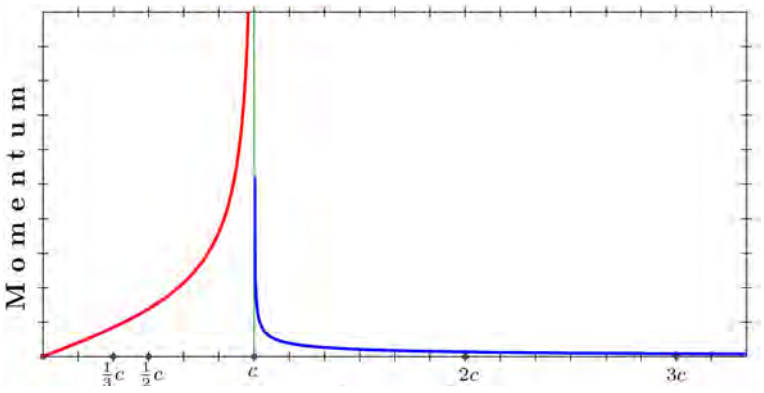


Figure 4. Dependence of relativistic momentum of the particle on speed v (red curve), the speed of light in a vacuum c (green line) and dependence of momentum of the primary particle on speed u; according to the primary particles hypothesis (blue curve).

$$W = \int_{x_1}^{x_2} F\mathrm{d}x = \int_{x_1}^{x_2} \frac{\mathrm{d}p_p}{\mathrm{d}t}\mathrm{d}x$$

and considering it is: $\mathrm{d}x = u\mathrm{d}t$, and that the rate of change of the momentum of the primary particle is

$$\frac{\mathrm{d}p_p}{\mathrm{d}t} = -\frac{m_p c^2}{2}\left(u^2 - c^2\right)^{-\frac{3}{2}} 2u\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{m_p c^2 u\dfrac{\mathrm{d}u}{\mathrm{d}t}}{\left(u^2 - c^2\right)^{\frac{3}{2}}}$$

$$W = -\int_{t_1}^{t_2} \frac{m_p c^2 u\dfrac{\mathrm{d}u}{\mathrm{d}t}}{\left(u^2 - c^2\right)^{\frac{3}{2}}} u\mathrm{d}t = -m_p c^2 \int_{u_p}^{u} \frac{u^2\mathrm{d}u}{\left(u^2 - c^2\right)^{\frac{3}{2}}}$$

$$= -m_p c^2 \left(\int_{u_p}^{u} \frac{\mathrm{d}u}{\left(u^2 - c^2\right)^{\frac{1}{2}}} + c^2 \int_{u_p}^{u} \frac{\mathrm{d}u}{\left(u^2 - c^2\right)^{\frac{3}{2}}} \right)$$

$$= m_p c^2 \left(\frac{u}{\sqrt{u^2 - c^2}} - \ln\left(u + \sqrt{u^2 - c^2}\right) \right)\Bigg|_{u_p}^{u},$$

we get the following value for work

$$W = m_p c^2 \left(\frac{u}{\sqrt{u^2 - c^2}} - \frac{u_p}{\sqrt{u_p^2 - c^2}} - \ln\frac{u + \sqrt{u^2 - c^2}}{u_p + \sqrt{u_p^2 - c^2}} \right).$$

For a slight decrease of the primary particle speed we see that

$$\frac{u + \sqrt{u^2 - c^2}}{u_p + \sqrt{u_p^2 - c^2}} \to 1, \ \textit{i.e.} \ \ln\frac{u + \sqrt{u^2 - c^2}}{u_p + \sqrt{u_p^2 - c^2}} \to 0; \ \text{and} \ \frac{u_p}{\sqrt{u_p^2 - c^2}} \to 1,$$

because it is $u_p \gg c$.

Finally, for work, we get the following

$$W = m_p c^2 \left(\frac{u}{\sqrt{u^2 - c^2}} - 1 \right) = m_p c^2 \left(\frac{1}{\sqrt{1 - \dfrac{c^2}{u^2}}} - 1 \right) = m_p c^2 \left(\xi - 1\right). \tag{47}$$

Work demonstrates how much the energy of the primary particle has changed, *i.e.* corresponds to its homeokinetic energy, as in Equation (43).

The relation between p_p and H is obtained by replacing (42) with (44) as

$$p_p = \frac{1}{c}\sqrt{H\left(H + 2E_p\right)}. \tag{48}$$

Also, from (40) and (44) we get

$$E_t = \sqrt{p_p^2 c^2 + m_p^2 c^4}, \tag{49}$$

which for $u \gg c$, *i.e.* $p_p^2 \ll m_p^2 c^2$ shows that $E_t = E_p = m_p c^2$.

However, in the case of $u \to c$, *i.e.* $p_p^2 \gg m_p^2 c^2$, we get that $E_t = p_p c$, or

$$p_p = \frac{E_t}{c}, \tag{50}$$

as we get from (46). The Equation (50) is valid for a certain high energy state of the primary particles and corresponds precisely to the relation between the momentum and the electromagnetic radiation energy which is obtained in the ultrarelativistic case for the Einstein equation $E = \sqrt{p^2c^2 + m^2c^4}$.

9. The Possibility of Interpreting the Creation of the Big Bang and Other Universes through the Collision of Primary Particles

The Big Bang represents the very event of creating of our spacetime, containing our matter and energy. The very moment in which this event took place, as well as the short period preceding Planck time, following the occurrence of this event, are still beyond the reach of science. Namely, Einstein's general theory of relativity does not work for systems smaller than Planck length, nor for the events lasting shorter than Planck time. By expectations of modern physics, this limit should be dismantled by the quantum theory of gravity, hence, this initial period of our Universe is referred to as "the quantum gravity era". We expect this period of the Universe to be explained by the primary particle hypothesis.

According to hypothesis, primary particles moving at speeds $u \gg c$ are in their flat spacetime. Simultaneously, their total energy (41) and momentum (45) are small. However, possibility exists that during collision of such particles, which would cause the speed of some primary particle to $u \to c$, an enormous amount energy is released from the point of collision according to the relation (41), *i.e.* the Big Bang would ensue. Because of the lower border speed of the primary particle at the Big Bang thus attained, the speed of energy-mass transfer in our Universe would have the value *c*, which is the same from the moment of the initial singularity. The energy thus released would be the very energy generated in the Big Bang in accordance with the conservation laws in physics, resulting in creation of our spacetime, containing our matter and energy.

Due to the huge number of possible lower border speeds *k*, to which the primary particles would slow down when creating other universes, their limitation of maximum speed of energy-mass transfer would be equal to instead of *c*, to some different *k*. It is clear that this would differentiate the values of other fundamental physical constants in different universes, as well. All relations stated in the paper would be valid in other universes as well, provided *c* is replaced with *k* from another universe.

10. Result

We assumed the existence of primary particles moving within their flat spacetime at a speed *u*, greater than the maximum speed of energy-mass transfer *k* in various universes, which, at the same time, represents their lower border speed. In addition to two postulates of the special theory of relativity, we introduced the

third postulate as well, according to which the values of the speeds of these particles are independent of the choice of inertial frames of reference from which they are being observed. In a thought experiment, we realize that dilation of time is the greatest while these particles are moving at speeds close to lower border speed, *i.e.* the speed of light c within our Universe, and get a ξ factor that corresponds to Lorentz γ factor. The same value of ξ factor was obtained via transformations of space-time coordinates for primary particles. We have shown that these transformations are Lorentz transformations. Based on the ξ factor, we have developed the dynamics of particles moving faster than light. We have proved that the assumed properties of these particles satisfy the fundamental laws of physics, which we have postulated. We recognized that the explanation of the Big Bang could be made from the standpoint of primary particles speed reduction during their mutual collision, during which the energy of the Big Bang would be released, simultaneously limiting the maximum speed of energy-mass transfer in our Universe to speed reached in that collision, c. The creation of other universes could be explained in a similar manner. Hence, the limitation to maximum speed of energy-mass transfer within them, as well as reached speed reduction of primary particles during big bangs in which those universes were created, would instead of c have various values k.

11. The Need for Development of a Primary Particle Hypothesis and Their Experimental Proof

It is clear that in this paper we have only made a logical assumption about some properties of primary particles. Thus, we believe their further study may lead to a major shift in physics, as well as our philosophical view of the world.

Based on the described properties of primary particles, we expect their indirect experimental proof to be possible through successful explanation of the Big Bang, via further scientific work on the hypothesis of primary particles.

Conflicts of Interest

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

References

[1] Morris, D. (2016) The Special Theory of Relativity An Introduction: Essentials of Physics, Dulles, Virginia.

[2] Petkov, V. (2010) Minkowski Spacetime: A Hundred Years Later. Springer, New York. https://doi.org/10.1007/978-90-481-3475-5

Journal of Modern Physics, 2019, 10, 1548-1565
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

5D Model Theory for the Creating of Life Forms

K. W. Wong[1], Peter C. W. Fung[2], W. K. Chow[3]

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

How to cite this paper: Wong, K.W., Fung, P.C.W. and Chow, W.K. (2019) 5D Model Theory for the Creating of Life Forms. *Journal of Modern Physics*, **10**, 1548-1565.
https://doi.org/10.4236/jmp.2019.1013103

Received: October 23, 2019
Accepted: November 17, 2019
Published: November 20, 2019

Abstract

Based on the Fermat's Last Theorem and the Po, P1 projections from the 4th space coordinate to the time variable for Po and to the remaining 3D space variables for P1, the carbon 12 nucleus is shown explicitly as given by the hard-sphere dense packing model that also satisfies the Gell-Mann standard model. It is through these that C12 is a vital element in all biomaterials, and all proteins as well as the Nitrogenous bases in DNAs, are of hexagon geometric structures. Furthermore, the unique presence of a 3D × 1D space void within the C12 nucleus provides for the monopole Boson field tunneling to occur, giving rise to the enormous variety spectra in the DNA of life forms. In addition, on the surface of the bio cells, the carbon valence band p electron excitation into the empty conduction band separated by a bandgap G, can result in HTC Excitonic induced superconductivity binding gaps from the Excitonic spectra, which match part of those of the DNA and thus produce the self-grow mechanism of numerous different cells in a life form.

Keywords

5D Projection, Nitrogenous Bases, DNA Spectra, Cell Growth, Genome, Protein Effect, Excitonic Superconductivity in Cells, Fractal Hexagon Carbon Representation

1. Introduction

We had forwarded a suggestion that "Life" can be modeled by the topological mappings from the 5D homogeneous manifold [1]. It was however only a hypothesis as in that reference neither the Perelman mappings [2] and the explicit Maxwell monopole potentials [3] were derived and incorporated into the 5D theory until recently [4]. The mappings of the homogeneous 5D manifold onto the 4D Lorentz manifold together with the imposing of the uncertainty principle with gauge invariance not only produce the Gell-Mann standard model because of the 4D space reduction projections giving rise to the Semi-simple Compact

Lie groups, SU(3) and SU(2) [5] giving the weak lepton model [6]. Furthermore, it also was successful in producing the models for the creation of galaxies with an extremely hot circular core, that would result in the presence of two gamma ray bubbles on each side of the core, while within the galactic plane there are the stars and planets together with their properties of self-rotation and the present of dipolar magnetic field [7] [8]. In these processes charges and masses are created by splitting the energy value of the Diagonal Long Range Order monopole Bosons along the Fermat's radial, which are originally in the Bose-Einstein condensed state, and hence giving rise to the concept of Temperature via the introduction of the grand canonical ensemble, yet retaining the Lie group symmetry as a fractal hexagon geometry of the mass structures, which are created by the space dimension reduction projections. It is therefore obvious that there remains the fractal representation on the extended low energy range or longer wavelength domain of the monopoles not responsible for the initial conversion into integer charges and masses should also exist for the 5D manifold monopoles. But since such Bose states must still be projected into geometrical symmetries that are technically the same as that given by the Lie groups. A simple illustration is the hard-sphere dense packing model for the carbon 12 nucleus [9]. Therefore, we see after applying the gauge constrain on the protons and neutrons in the nucleus that the topological symmetry is the formation of 3 hexagons and 6 equilateral triangles [1]. Hence in order that the projection action can be applied also to the very long wavelength region of the monopoles, a breaking of the 3D spherical space symmetry given for the carbon nucleus into 2D × 1D must be imposed. It is this further space symmetry breaking that leads us to the basic formation of the 2D biological Nitrogenous bases, that is realized in a correspondingly lower temperature domain from the canonical ensemble. However, due to our need to maintain the gauge invariance on the DLRO monopole Boson by completing it into a closed loop, such 2D hexagon-shaped Nitrogenous bases must be linked along the 1D via quantum tunneling. This results in the creation of RNAs and with the implications of an endpoint reflection into a parallel RNA and therefore completing a DNA [1], the most essential component to the building of "Life". The diversity of such DNAs formed is infinite, as the thermal parameters are varied, not to mention the infinite possible variations of the gap magnitudes between the different base layers, hence the infinite diversity of life forms. It is this most complex geometrical projected results, which we proposed as responsible for "Life" that we like to present in the current paper as the physics of "Life". In the present paper, we shall utilize these mathematical results and separate our discussion into the following separate sections linking DNAs to genome and the creation of proteins and cells, etc., and thus life forms including us humans.

2. The Mapping of the Magnetic Monopole Field within an Enclosing Time Frozen Matter Poincare Sphere

The Maxwell monopole fields are the quantum field result of the 5D homoge-

neous quadratic space-time operator. It is a Diagonal Long Range Order DLRO product of two opposite "e" charge massless spinors, that is along the Fermat's amplitude "r" [10]. Thus its magnitude is given by "2ec". Since it is non-dependent on the eigenenergy value of the monopole Boson it must be represented in the grand canonical ensemble and thereby in the Bose-Einstein condensed state as we mentioned in the introduction. Such a state implies that the magnetic monopoles give raise to a vacuum in 4D space-time that is filled with this zero net charge Boson field, a quantum picture that is consistent with the Higg's vacuum [11]. We refer our readers to ref. [11] for details. Hence, should the 5D manifold be time frozen, the net 0 charge magnetic monopole 2ec along "r" must be mapped onto the inner surface of the spherical void instead of being broken into 2 separate 0 and 2e Boson fields circulating within the 3D spherical void core through Po, the space to time projection that resulted in SU(2) for the -e leptons and the SU(3) via the space to space conformal projection P1, as discussed in detail in ref. [1], giving us the Gell-Mann standard model, which are resulted from the monopole energy, and thereby resulting in the rotation angular momentum state <2hvt(0)> along z, where <> represents the canonical ensemble averaged, and a separate magnetic dipole B'_z, as proportion to $M.r = <2ec^2.t(0)>$. Both such results were discussed at length by us as illustration on the self rotation and dipolar magnetic fields in stars and planets [7] [8]. We again refer our readers to these two earlier publications. What we need to emphasize is matter created via Po and P1 are from monopole energies exceeding the rest mass energy of the electron and the bare quarks by the choosing of a different chemical potential value m(e) the electron rest mass for the separated negative e charge, and m(q), the quark rest mass for that of the positive e charge. The detail obtained through force balancing due to the Perelman mappings and the masses from Po and P1 are also recently published [12]. Again we refer our readers to that paper. However, the lowest quantum Poincare sphere resulted from P1, gives us the carbon 12 nucleus [13] was not carefully analyzed. In fact it was suggested in ref. [1] by making sure that the 6 protons and 6 neutrons on the C12 nucleus shell form 3 hexagons and 6 equilateral triangles, with each point triple degenerates, hence making them in the lowest Standard model that of either a proton or a neutron and makes the nucleus C12 the lowest symmetric Poincare sphere, with a spherical time frozen void core [**Figure 1**], absence of the monopole field. However within the low energy frequency region this DLRO field must remain intact, but it remains in DLRO when time is frozen, unless it can quantum tunnel from C to C and completes a closed loop, as required by gauge invariance. It is this realizable gauge state that introduces the fractal representation with the replacement of the proton and neutron in the C12 nucleus with a C12 and a Nitrogen atom sphere without a void core, such that we may generate 3 pairs of cancelling + and −Bz' along "r". With 60 degree in between each z'(j). Because of charge equipotential Theorem, such quantum Bz' states will always cancel each other under perfect symmetry,

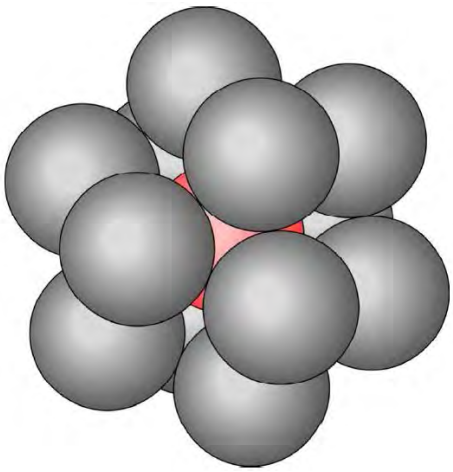


Figure 1. C12 hardsphere dense packing model (Wong *et al.* 2014, ref. [1]).

unless we can divide them and map the 3 hexagons produced by P1 onto a flat 2D plane. It is this topological surgery that creates 3 totally independent Nitrogenous bases, with 3 neighboring C's, and 2 N's sandwich a C [see **Figure 2**]. Because the 4th C between the 2Ns only serves to create a bone structure linking stacks of such Nitrogenous bases, forming RNA structure, it cannot be considered as also a fractal representation of the proton by C12. Hence by recognizing the M DLRO state must remain perpendicular to the hexagon, this monopole Boson field must tunnel from one layer of the Nitrogenous base to the next, provided by wave-lengths exceeding the space gap between the two neighboring layers, which could also contain non-Nitrogenous C layers, so that at the ends it must be able to be reflected onto a parallel RNA stack, thus completing a mathematically required closed loop and thus preserves the gauge invariance. For the two RNAs within the DNA end points reflection an extra reflective boundary condition hence a 4th Nitrogenous base representing the fractal of Po mapping must be introduced to make the basic set in all DNAs, 4 bases instead of 3 as depicted by fractal geometry. With such rather long wave-length spectra, because of the large number of non repeated choices = 8! Of stacking of the parity pair Nitrogenous bases layers in the DNA, the equivalent temperature energy is then in the so call life suitable temperature range. In order that the stacking can complete a close loop by requiring that the two ends of the double RNA, that makes up the DNA be always those of the Thymine base, the number of non repeat stacking is given by the 4 bases with their mirror representation, as they are not necessarily parity invariant. While the gene frequency is dictated by a closed loop of the M, the DLRO Boson that quantum tunnels from the 3 adjoining C's in a base to the next 3 C's in the next nearest neighbors bases layers. Hence the genome number must be given by $3 \times 8! \times 2/7 = 34{,}560$. The number 3 comes from the 3 adjoining C's within each base. Since each quantum closed loop of M must pass through all 8 bases states, the number of different possible path or

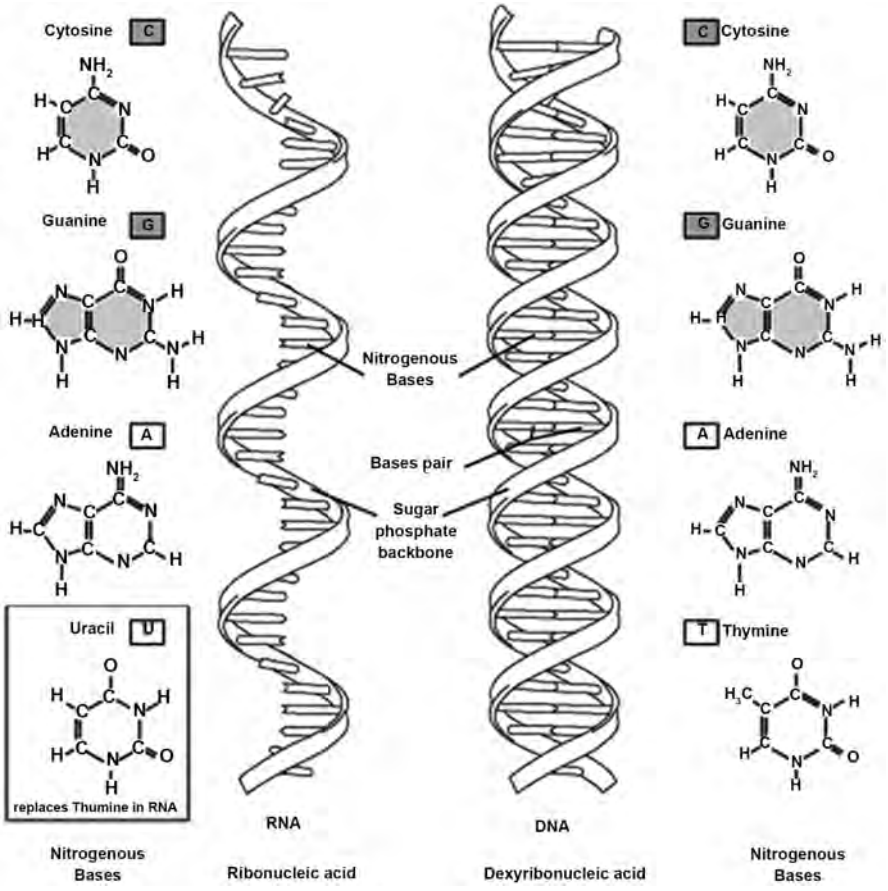


Figure 2. Images of DNA and RNA [Wong *et al.* 2014].

dering is 8! However the 2 end cap Thymine bases are parity invariant, and must be separated by 6 non Thymine bases in between within the closed loop, thus the number of distinct choices is multiplied by 2/7. The 2 comes from the two Thymine bases within. The 1/7 is due to the fact that once one Thymine cap is chosen then the remaining Thymine must be fixed and hence is only 1/7 coming from the 7 bases locations left to choose. It is interesting to observe that this genome number appears to be universal to life forms, it's monopole M eigen-frequencies depend on the loop circumference, which must depend on the base to base separation gaps, as well as the DNA twisting, and thereby depend on the physical length of the DNA, that is vital to the creation as well as serves as a signature of an individual life form. Hence comparing the DNA spectra is then applicable as a verification of the ancestry of a life, because the quantum frequency spectra on the DNA loop is quite unique, with infinite possible separating gaps variations, hence no two individuals are 100% exactly identical. The key we emphasize is the quantum frequency spectra range is dictated by the stacking geometry, which not only included the multi-non-Nitrogenous in-between layers. Thus as averaged over a thermal canonical ensemble has a temperature T spread. It is such a T spread that dictates whether the DNA generated life can exist in its thermal environment. For example, sea lives due to water freezing at 0

C. would have a narrower T distribution for its DNA then human that can survive a much wider temperature range. The wider the T distribution for the DNA spectra, the more complex of a life form it tends to produce.

Only roughly 30,000 genome is found in human.

3. The Role Played by the DNA Spectra and the Natural Growth of Biological Cells

All biological tissues contain proteins. All proteins are basically 2D structures of molecular hexagons containing carbons. It according to 5D theory are the result of Perelmann entropy mapping, combined with the chemical binding of other elements within the thermal environment, similar to the formation of galaxies, except it has no center void core, due to the dimensions of the macro-hexagons formed, hence it also does not contain self rotation and a magnetic field. Since they are fundamentally 2D, they are very flexible and hence the elements within do not form a rigid point group symmetry. That is the electron orbitals within these protein structures do not result in a band structure. The forms of proteins are thereby purely dependent on the availability of elements within its thermal environment. Because of the fractal picture as we mentioned earlier derived from the P1 mapping of homogeneous 5D manifold, which gives us the C12 nucleus, and the basic hexagon structure, all tissues are formed in strings as can be visible from the topology of the next order Poincare sphere symmetry; namely the Bucky ball, where the hexagons are linked with isolated pentagons in between [see **Figure 3**]. Making biological tissues into string structure fibers. Like crystals it can be broken by temperature changes. Yet these proteins are vital to the formation of cells. Cells on the other hand are basically spherical, with dangling bounds attached to the surface, and always contain DNAs within it. Because of the spherical nature, the electron orbitals on the surface of the cell, necessarily obey periodicity and thereby form a band structure. When proteins are rapped around a DNA, the spectra of the DNA affects the transition of the cell's surface electronic states into an ODLRO, such that its binding energy matches precisely of the eigenvalues in the DNA's DLRO gauge invariant monopole Boson state. The details of this mechanism will be discussed in the next section. Since there are a large number of DNA spectra eigenvalues, there are correspondingly a large number of different cells possible. The introduction of an abnormal protein into the bio environment, can have a detrimental effect on the life itself, as it creates abnormal cells, or inhibits the normal cell formation. Hence a normal protein belonging to a specific life form, like a plant that is absent in animals can be toxic, which includes us human. On the other hand the reverse are beneficial to the animal when consumed. In fact it is the principle behind the benefits in herbal medicines.

4. The Key Role Play by Carbon 12 in Bio Cells

There are numerous studies and publications on many different areas in cell biology. Its functions, compositions and reproduction: Including genome and

Figure 3. The 20 hexagons are in white, while the 12 isolated pentagons are in black.

chromosomes. To ignore these progresses is a huge mistake. However, there has never been any study on the creation of life from basic principles of space-time mapping and fundamental quantum fields. It is this note's intension to initiate such a study based on the fact that all known life forms are of carbon composition.

Carbon 12 is a very unique element based on the 5D theory [1]. Its nucleus has 6 protons and 6 neutrons, forming in terms of a hard sphere model [10] into 3 hexagons and 6 equilateral triangles on a shell, enclosing a spherical void core. From the 6 triangles, it is obvious they must contain either 2 protons with 1 neutron or 2 neutrons with one proton at the corners. Thereby the charge distribution cannot be uniform in alternate proton to neutron throughout. In fact if we consider the probability distribution on the shell of finding a charge it would be either 2/3 or 1/3. These fractions are actually in agreement with the P1 space to space projection model from 5D homogeneous space-time with uncertainty principle, where only e and −e opposite momentum pairs of massless spinors along the Fermat's amplitude can exist, such that the +e of the massless spinor is broken into fractional charges and gain a mass becoming quarks in the 4D Lorentz space time L, so that the 5D is mapped into SU(3) × L × 1D [6]. From the Fermat's Last Theorem, the homogeneous space-time N dimensional manifold are expressed as the amplitude r, with angle $0 < \varphi < 2\pi$ for 2D space, $0 < \theta < \pi$, $0 < \varphi < 2\pi$ for 3D space, and $0 < \theta < \pi$, $0 < \varphi < 2\pi$, $0 < \varepsilon < 4\pi$ for 4D space. Thus under the P1 projection from 4D space to 2D space, in terms of hard spheres, the 3 hexagons due to ε and 6 equilateral triangles geometry resulted on the spherical surface. But because of the equipotential charge distribution Theorem, the superposition of the 2 equilateral triangle charge representations choices must be imposed, such that the +e charge is broken into the set of SU(3) generators: +2/3e, +2/3e and −1/3e, which is equivalent to the realization of gauge invariance preserving unit +e. And at the vertex points of each of the 3 hexagons representing the ε phase, which has a total of 4π internal net angle, thus it is a

geometric realization that represents the 4th space dimension as according to the Fermat's amplitude and phase angle representation.

While C12 nucleus is composed of 18 up and down quarks each, divided into 3 hexagons, and making every vertex point containing 3 quarks that satisfy gauge invariance, such that the C12 nucleus is actually a quantum quark soup. Thus on the C12 shell, we have exactly 6 protons and 6 neutrons in agreement to the nuclear shell model. So that leaving only within the C12 void, the radial outward net 0 charge and massless Bose magnetic monopole fields M = +,−2ec can remain to exist [3]. While the projection Po from the 4th space coordinate onto time must produce the 6 SU(2) −e electrons so as to maintain charge neutrality of the 5D manifold from which C12 is created by space dimension reduction projection. Hence in the atom C12 gives the 6 electrons revolving around the nucleus. It should be mentioned that if a spherical 3D × 1D time frozen void exists within the earth crest, and if fractal representation is valid, not considering charge, a very unexpected rock formation of hexagonal shaped columns composing of mainly 2 elements within the mass of granite rock due to P1 and Po due to 4th space dimension reduction mapping [4] could be formed even from volcanic eruptions after the molten lava cools. This was of course found throughout the earth's surface [see geological data in south China, HK history museum]. In yet another unrelated natural compound, is the Lithium carbonate, LiCO, a key ingredient for the Lithium battery is based on the relatively low photon energy ionization of Lithium, and it is easily recharged, because of the storage of magnetic monopole tunneling energy within the carbon on top of the carbon hexagon loop structure due to the carbonate structural composition, that can also readily converts back to the ionization photon to restore the battery charge. Such Lithium carbonate is found natural in the Brien salt field, which reviews on the surface the hexagonal structure, a fractal representation of the Sodium carbonate [see National Geographic, Feb. 2019]. Implying the topological importance of the hexagon structure due to the conformal space to space projection is valid in all energy range, and not just for the standard model of hadrons.

The Carbon element is a covalent atom, in fact the most important covalent element in nature. To understand how, let's go back to the basic Bohr atomic model. The eigen values of these 6 atomic electrons are according to Bohr's model, given by - ZR/n^2, where n is a positive integer number, R is the Rydberg constant and Z the n level effective charge number. Thus for n = 1, level Z = 6, and for n = 2, Z = 4, as it is reduced by the 2 n = 1 inner s orbital level electrons shielding. Therefore for C with 6 electrons, its 4 outermost orbital level must be given by n = 2, with 2 s states and 2 p states. Since the p orbitals are given by 2j + 1, where j = l + s, it is 1/2 full, while s is filled. This means carbon's atomic outermost p and s orbital electrons can be shared with its molecular neighbors in binding provided from its valence states by either filling the remainder 2 unoccupied quantum p angular momentum states or by giving its 2 p or 2 s to the binding neighbors as energetically allowed. This 1/2 filled p or empty s after binding is reflected as an example by the photo absorption ob-

served in graphite, because of the continuous frequency spectrum allowed in a half filled VB, making graphite appears as black. Since the graphite structure is a stacking of weakly bond 2D C square lattice, which means it obeys space symmetry of 1D × 1D geometry. And since 1D does not allow for angular eigenvalues, hence in graphite the unoccupied VB band must be the s states.

Because basically the s hole state is flat, it has very large effective mass, thus making graphite a very poor hole conductor. However, K inoculated graphite is a superconductor with a very low Tc temperature. This LTC superconductivity mechanism of K inoculated graphite has never been properly explained in the literature. It is actually easy to explain the mechanism, as that inserted K+ ion must give away its s electron to the C atoms in the inoculated graphite filling a s hole between the layers as given in the VB band, making it 2D × 1D like, thus slightly changes the hole density composition, by exchanging s and creating a small fraction of p density in the residual VB hole density. But because this density of the p hole created due to inoculation is so small, the 2D like Exciton that can exist in the band gap G by the electric dipolar exciting from VB an electron into the empty CB, caused by the VB p hole coupling is also weak, thereby according to the EEM [Excitonic Enhancement Mechanism] mechanism, [14] produces [the EEM theory was introduced for the ceramic HTC materials, and has been extensively investigated both theoretically and experimentally [15]-[21] An example list of such works is given in our references. Readers in particular from the bio sciences, interested are advised to read them] a quantitatively low Tc. Hence such a K inoculate dgraphite superconductor, despite the low Tc value, actually is due to the same EEM mechanism that produced the ceramic HTC, and not that of phonon induced BCS mechanism in superconducting metals [22].

In fact since in general molecular bio structures of C, the C p orbital electrons with lattice geometry consisted of hexagons it can equally share either as a negative valency state in molecular binding, or a positive valency and accept an extra electron to fill its p angular momentum shell irrespective to the rest of the nearest neighbors geometry. In such more general molecular bio chemical rigid structure, the C's in the structure would naturally produce a partially filled p VB structure, when the bio carbon forms a solid structure with translational unit cell symmetry. But unlike that in graphite, because of the VB half filled p states will not have large effective masses. In fact because by changing the 2D C net from square to hexagon, changes the number of nearest neighbors from 4 in square net to 3 in hexagon net, leaving the VB with at least one p hole, which can have relatively low effective mass. All proteins that make up bio cells have natural hexagon C structures, except free proteins are not rigid.

Therefore in free proteins, the electron periodic boundary condition needed for band theory is not met, and hence in free proteins we cannot produce defined band structures. On the other hand, all bio cells are basically 3D spheres and contain carbons from the proteins coating in its molecular surface structure, which can be made rigid and enclosing a DNA within it. And in a fractal way, by

taking C as replacing the proton in the C12 structural form, because of the M state within DNA core, and the P1 that broke the 3D × 1D space symmetry, producing the hexagon structure shell for the cell, giving us the band structure for such cells, which are in a macro scale in the geometric shape of a sphere except also normally with different 1D like spike attachments. The effect of such attachments does not break the spherical surface periodicity on the electronic states from the C12 atoms, but like K inoculated graphite it would modify the VB p wave density and give a perturbative correction to the exciton spectrum formed within the band gap G [23]. Such minor spectrum changes are vital to the perfect matching to the DNA in the core spectra, and thus able to induce the electrons ODLRO [Off-Diagonal Long Range Order] in the specific cell, and with it the Tc value. It is such spikes that can also give the cells different functions as according to the DNA and its ability to be linked and build an organ structure. A change in Tc upward can result in inducing a cancerous growth of such a cell. The specific organ forming requires specific boundary conditions imposed not just from the DNA spectrum. Such complex boundary usually includes the chromosomes, etc. Are beyond the scope of this simple note. An example is the form of the free white blood immune defensive cell. It does not come from the VB p holes, very likely such memory comes from the semi-conductor portion on the attaching surface, similar to the LCD screen of a TV? Thus such surface semiconductor circuit memory works like a computer chip, and produces the life form's memories. It is all throughout the body of the life form, not just in the brain. The body chromosomes are examples. Although the brain cells are unique, in that its circuit accepts external programming commands, and not just play the role of the organ and shape formation. Its purpose is mainly to act as a conducting net for the stored commends in the brain provided by programmed storage key. Thereby the brain cells are generally dense packed, such that super currents can propagate on its surface similar to the granular dense pack HTC [24]. It is this physical structure that makes neuro cells rejuvenation rather rare, and its death enhanced by usage, thus increases with aging [25]. Most other bio cells, are free floating in body fluids, and it is replaced frequently, as the DNA responsible is enclosed by the specific protein that forms it. While the dead cell, breaks apart, and the protein cover is oxidized and removed from the body as CO_2 and other gases emission. Because, there are few blood vessels throughout the brain, the REST protein released from death brain cells usually will accumulate and covers the brain structure, thus blocking the super current flow. It is this current blockage that leads to human neuro deceases such as Alzheimer and Parkinson. If our explanation of the cause of such deceases is valid, then we can suggest a method of control, but not a cure, if we can somehow increase the oxidation of the REST protein? Since we have no way to increasing the number of blood vessels around the brain, the only way to increase oxygen then is to enhance the blood flow. As is well known, that consuming alcohol leads to intoxication due to alcohol tends to flow to the brain. Hence, if the alcohol contains a strong oxidation agent, such as red grape wine,

then perhaps by controlled wine drinking might keep the REST protein quantity in check? Of course such method must be tested, and success might be very limited! Thus the brain cells are susceptible to blockages due to surface coating of extra proteins, not necessarily becoming cancerous and yet leading to memory loss. Nonetheless all the cells surfaces are fundamentally spherical, and all the electronic states on its surface can be periodic as the basic chemical molecules on it always satisfy 2D periodic boundary, even if it is only a single unit of a cell, like for example, a Bucky ball with just certain fixed points on the surface being that of C. In short, the electron structure from them would satisfy the boundary required for band structure calculation. As such, and with the covariant character of the carbon orbitals in such spherical chemical molecules, would necessarily give rise to a partially filled p state VB, together with an empty CB, which will have from n = 3 atomic level d orbitals, very much like in the HTC cuprates. Since the bio cell normally will not combine with a free foreign atom, unless the free atom attachment can bring down the binding energy in excess of the band gap G between the cell's CB and VB, otherwise such a cell will not be stable when submerged in a fluid containing free ions, in bio terms, known as free radicals, when there is no positive band gap energy G, separating CB with VB of the cell band structure. Such a band structure with positive G resembles that of HTC cuprates, both band structures consisted of partially VB, with 2D like p hole conductivity in the normal phase, except in HTC the anisotropic conductivity is the result of layers of CuO planes stacking, while in bio-cells it is mainly due to C's on the spherical surface [23]. The molecular orbital for CuO in the bands gives a partial filled p and d alternating states, where p comes from the oxygen ion, unlike in graphite, thus with relatively small effective masses [14]. In fact the experimental result on Tc dependence on oxygen deficiency in YBCO supports our EEM theory interpretation for HTC mechanism.

Since the excitation of an electron from VB to the empty CB would form an exciton state, within the positive bandgap G, because such atomic like 2D exciton due to the cell's periodic boundary, if formed is equally periodic on the structural surface which as required for formation of band structure, it can actually couple to any moving phole by the electrical dipole induced by the Photon transition excitation of the orbital states caused by the hole charge motion. It is this electrical dipole coupling between the periodic exciton and the mobile holes that will produce EEM and thereby creating the superconducting phase transition for temperature below its superconductivity critical Tc [15]. Since Tcdepends on the Excitonic dipolar coupling, it therefore is indirectly dependent on the exciton's quantum orbital states. That is the set of reduced charge quantum orbital excitation transitions. Hence when such photonics spectrum photons are made available from the enclosed DNA, the cell would naturally form when the protein molecules that it is made up of are available within the thermal bath, so that the bio structure can minimize its carrier ground state binding gap energy matching part of the DNA spectra according to that of the enclosed DNA.

The lower the set of photo energy transitions, the easier it will be for the p

holes in the cell ODLRO condensation to occur, leading to the natural growth of the cell. In fact, the closer the exciton level is to CB, the higher would be the specific cell's Tc [23]. Meaning, the higher Tc value as compared to the thermal bath temperature T, and the more stable it is for the cell. None the less the EEM induced Tc has an upper limit, due to the optimum half filled VB p state, and its effective lowest possible mass, as due to the unit cell's size. Hence all life forms can only exist within this limited upper Tc range, normally below the boiling point of water. While the lower temperature limit is bounded by water freezing point, as water provides the fluid for moving the free proteins around the bio structure. Heat generation within a bio system, usually caused by Oxygen and other chemical reactions between the body fluid ions, such as free ions from proteins, which are C hexagon structures, but flexible without fixed bond lengths and can be transformed by wrapping onto a spherical structure cell, thus changes it, plus sugars which are just C, silicate, and amino acid etc. that can make it rigid, and could also destroy the original cell structure thus breaking the p holes ODLRO phase, hence leading to the natural cell death. In fact, such stiffening of the shell of the specific protein is vital to the cell normal growth and death. To that water plays as a facilitator, an indispensable role in the chemistry. Chinese traditional medicines rely on identifying the specific protein associated with the Heath of the specific cells that is responsible for the specific organ operation, thus hoping by such direct supply of the needed protein can restore the organ's proper function. Although overabundance of such free radicals within the body fluids, will cause change in the normal cell regrowth cycle, and will cause health problems, like diabetes, Alzheimer and Parkinson's decease when the protein wraps over a brain cell is improbably stiffened with excessive silicate, causing memory loss and abnormal aging. However, the replenishment of the body fluid, with proteins coming from food intake is an essential component to the regrow of the shell of a cell around the DNA, and is equally also important to the body organs that must come from the build-up of many different cells via also providing the inter linkages, including the rigid silicate bone structures precisely maintained within such complex thermal cycles that operates in the bio system to create a healthy life.

Should the thermal bath temperature rise beyond the cell's superconducting Tc, due to inflammation, normal cell growth would be prohibited, and only abnormal cancerous ones survive. Hence when the body suffers from sickness, it is usually accompanied with at least a localized rise in temperature, the occurrence of inflammation there inside the body. Actually, the thermal condition and the availability of the proteins and chemicals drives the species evolution, or the minor topological mapping of cells, etc. towards its adaptation to survival. In order to apply our Tc model principle to the cause and potential cures of the body, all the exact normal cell structure and thereby its band structures must be obtained, so that the proper Excitonic spectra that can be formed within the corresponding band gap must be derived and matched to the body's DNA spectra. RNAs and DNAs of bio systems are made from 4 basic Nitrogenous bases,

stacked and formed by back bones into rigid structures. These bases are all composed of hexagons with 3 C carbons, with 1 C carbon isolated from the others with an N nitrogen in between [see figures]. Since in our 5D model, it is the Po and P1 projection operations that lead to converting the massless charged spinors into SU(2) gauge conserved leptons and SU(3) with fractional charged quarks that create matter, thereby through the Perelman-entropy mapping giving us the Nitrogenous base structures, are divided into 3 + 1 distinct attachments as shown in the figures. The 1 represents the end base cap, responsible for reflecting the M field from one RNA branch into the parallel one, while the 3 together with its opposite sides represent the 6 SU(3) generators, since each C contain a time frozen 3D × 1D void, where the magnetic monopole Boson State M = 2ec can exist if it can tunnel from C to C and completes a closed loop, conserving gauge invariance by the stacking of two parallel chains of RNAs, with an end cap base to satisfy the formation of a DNA.

5. Summary on the 5D Projection Processes

Because the creation of 'Life' is the most challenging question, yet we believe it can be explained based on the assumption that the universe started with the creation of a single homogeneous 5D space-time manifold, which satisfies the Fermat's Last Theorem [10], together with the uncertainty principle on measurement, thus leading to the Big Bang theory, it is best to give a summary to the steps that leads to all creations: 1) The quantum field solutions of the 5D homogeneous manifold are those of the Electro-magnetic 4 vector potentials, plus the extra magnetic monopole potential as suggested by Maxwell. This Maxwell monopole magnitude is a Bose field of 2ec from the product of e and −e massless spinors, where e is the coupling between the Electro-magnetic potentials and the massless spinor, thus must obey the gauge transformation. 2) Since the Fermat's sum is only valid for positively increasing time, these vector potentials are not time reversal invariant. And it also means, parity symmetry need not be rigorously obeyed. 3) The magnetic monopole Boson is energy independent, and is of Diagonal Long Range Order, therefore, it must be in the Bose-Einstein condensed state. But to become that it must be in a Grand Canonical Ensemble, thus making up a chemical potential value in exact cancellation to the Boson's energy. 4) The fixing and choosing of a discrete set of chemical potentials break the Boson ground state, similar to the Higg's vacuum theory [11] [26] by breaking the continuous chemical potential of the Grand Canonical Ensemble into a discrete set. 5) The creation of mass matter, is then identified with this chemical potential set, and is an irreversible process, equivalent to the space reduction projection, and the Perelman mappings. Hence one ties the irreversible mapping for the creation of mass to the creating of Lorentz 4D manifolds out of the original single homogeneous 5D manifold. As well as identifying the breaking of the net charge 0, monopole Boson into separate +e and −e massive spinors [1]. Detail on the symmetry breaking had been discussed in earlier publications and will not be

repeated here. Nonetheless it clearly divided the energy range of the monopole Bosons, into a range when its energy is greater than the rest mass of the electron, the lightest massive particle due to the projection to the lower energies. 6) Due to the rest mass differences between the +e and −e massive spinors generated by the Boson ground state excitation, the resulting atomic elements, which composed of protons, neutrons and electrons no longer maintain charge and parity symmetry [27]. 7) Since the breaking of the continuous energy range for the monopoles due to uncertainty principle also implies a breaking of the continuous 4D homogeneous space, thus would result in the separate formation of doughnut disc like Lorentz manifolds, which in turn divided the universe into disjoint Lorentz manifolds, with multiple energy source composed of the distribution of monopole Bosons, from which both +e and −e masses can be created. Below the rest electron masse region of the Monopole Bosons, it is obviously of longer wave-lengths. Thus under the Perelman entropy mapping [3], Poincare spheres can be obtained, with a resulting time frozen void core, and the forming of masses into three hexagons shape due to SU(3). While the void core dimension restricts the existence of the monopole Boson field. 8) The breaking of the spherical masses on the shell further splits these 3 symmetric hexagons into 3 2D separate hexagons, separated along the vertical axis of the 2D plane. 9) In life forms, we are interested in the monopole state presence due to the void of carbon 12 nucleus, as all life forms are basically carbon compounds. Since the carbon 12 void is of nucleus dimension, the monopole Boson would have to have a short wave-length and does not exist normally, unless it can quantum tunnel from one C in one hexagon onto another C in the adjacent hexagon layer, thus creating the RNA, except with fixed end points, as represented by the Nitrogenous bases set: Cytosine, which comes from the fractal representation of the hexagon on the C12 nucleus surface [**Figure 2**], while Guanine, Adenine are from those on the Bucky Ball surface [**Figure 3**], due to the 2:1 ratio between numbers of pentagons in the Bucky ball to the equilateral triangles in C12, and the 1D end cap Uracil, such that the monopole Boson is a standing wave [**Figure 1**]. Hence this Boson state violates gauge invariance. 10. In order to retain gauge invariance for the Boson wave, we need to change the end cap hexagon into a reflecting cap, Thymine, such that the Boson state from one RNA reflects onto a parallel RNA, and becomes a closed loop forming a DNA 11. The non-repeat stacking of the bases then gives us the genome, with its unique spectra for each life form. 12. This unique quantized gauge solution set of the genome spectra for the monopole Bosons allows the computation on the thermal temperature T by using 1D bose distribution averaging within the life form's body, [note that 0 frequency state is always excluded, such the genome spectra does not B. E. condense] and in order that cells can be generated by the EEM mechanism for the partially VB p holes from the carbons in the protein shell the superconducting critical Tc must exceed T. Yet not too far as to affect the normal cell replacement cycle caused by oxidation, and other chemical reactions. This balancing of T criteria in

facts dictates the well being of the life.

6. Conclusions

Under the homogeneous 5D space-time quantum theory, the creation of matter from the orthogonal 4th space Bose-Einstein condensed magnetic monopole Boson fields, in the Grand Canonical Ensemble, due to space dimension reduction into Matter filled Lorentz 4D, with 5D voids, can be divided into different energy regions, and with high energies above electron rest mass, splitting the equal energy values of + and −e massless spinors with the different chemical potential values according to the electron rest mass and the quark rest mass for the different charges in the monopole field amplitude in the Grand Canonical ensembles leading to the realization of the Gell-Mann Quark model as well as the electro-weak lepton model, which we had previously published. It also can be compared to the breaking of the Higg's vacuum. Thereby, there remains the energy region below the electron rest mass, which can be shown to further split into 2D space matter geometry of hexagon structure when the remaining 3D homogeneous space is broken into 2D × 1D representation. This 2D space includes that of a spherical surface, and that of a cylindrical cross section representation which would result the lowest symmetry geometries, that of the nucleus C12 shell, that consisted of protons and neutrons in 3 hexagons and 6 equilateral triangles and for the cylindrical geometry 4 basic Nitrogenous bases, formed by carbons and nitrogens also in hexagon form as the result of fractal where the nucleons on C12 shell are replaced by the carbon and nitrogen atoms that are spaced along the 1D, thus creating both RNAs and DNAs. It is through these results and gauge invariance genome and chromosomes of life forms can be obtained. The distinct discrete set of such spectra coming from the Bose-Einstein condensed DLRO of the monopole Boson field would then induce the enclosing matter surface elements' CB electrons, or holes of unfilled VB to become superconducting ODLRO. Hence from the 5D quantum model, the Quantum Long Range Order continuity causes the DNA spectrum energy to match the ODLRO electron or hole pairing energy thus creats cell growth, and thereby "Life". In fact this space symmetry reduction is equivalent to the topological projections of Po and P1, discussed in the 5D unified field theory book [1].

As such the P1 created M tunneling states made C into the fractal representation as given by the 3 adjoining C's in the Nitrogenous bases of the SU(3) quark generators, while the Po was represented by the isolated C. Since Pogives raise to the SU(2) electron generator, that would be in the atomic orbital and responsible to the molecular binding, so this isolated C acts to hold the Nitrogenous bases stacking together, creating the back bone and hence must be responsible for the RNA/DNA twisting back bone required to make them rigid, and hence fixes the allowed M eigen spectrum. Due to the many possible Nitrogenous bases stacking order possible, together with the twisting as well as various layers to layer sepa-

ration gaps, the eigen spectra of M is extremely large, enabling it to match that required to form the exciton induced ODLRO gap in the cells. Hence, using suitable RNAi technology it is possible to correct and cure all unnatural cancerous growth, and cell inflammation caused by mismatching of the DNA spectrum with the cell ODLRO gap formation. Although this is a very tedious computational problem, as there are numerous different cells, which serves to maintain the proper functioning of different organs and senses in the life forms. Nonetheless it is an achievable task with our current 5G computer computational power. It is to this objective end, that we hope if our 5D model theory for life is correct, can help bring about a quantifiable medical revolution method in cure for many yet incurable illnesses? Should our model be correct, this enormous work would make it the greatest technological revolution in human history. The quantitative spectra matching between DNA and the bio exciton states would identify all the life forms functioning and its senses, including the brain's memories. As a result, there are numerous patents to be filed, and will involve collaborations between cell biologists, band structure computational physicists, computer engineers, programmers and RNAi medical and bio technicians. The results would produce advances in new technology from agriculture to human health. We, as pure scientists should lead this revolution, bringing it unconditionally to the world, instead of letting selfish politicians gain control, which could lead to human annihilating world war just so they can use the technology to satisfy their ambition in gaining power for world dominance!

Acknowledgements

We sincerely thank Ms. W. L. So for her help in typing and editing this paper. Simultaneously we thank Professor Dreschhoff and Professor Jungner for insightful discussions.

Conflicts of Interest

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

References

[1] Wong, K.W., Dreschhoff, G.A.M. and Jungner, H. (2014) The Five Dimension Space-Time Universe—A Creation and Grand Unified Field Theory Model. Scientific Research Publishing, Wuhan.

[2] Diklan, J. (2019) You Are Disturbing Me—I'm Picking Mushrooms! Academia.edu. Jan. 2019.

[3] Wong, K.W., Dreschhoff, G., Jungner, H., Fung, P.C.W. and Chow, W.K. (2018) The Magnetic Monopole in 5D Homogenous Space-Time. *Physics Essays*, **31**, 493-495. https://doi.org/10.4006/0836-1398-31.4.493

[4] Wong, K.W., Fung, P.C.W. and Chow, W.K. (2019) A Quantum Representation of the Homogeneous 5D Manifold and the Perelman Mappings of 5D onto Non-Homogeneous Lorentz 4D Manifolds. *Journal of Modern Physics*, **10**, 557-575.

https://doi.org/10.4236/jmp.2019.105039

[5] Wong, K.W., Dreschhoff, G. and Jungner, H. (2012) On Neutron Oscillation and Predicting the 125 GeV. Two Photon Emission State from p-p Collision Based on the 5D Homogeneous Space-Time Projection Model. *Journal of Modern Physics*, **3**, 1450-1457. https://doi.org/10.4236/jmp.2012.310179

[6] Gell-Mann, M. (1964) Nonleptonic Weak Decays and the Eightfold Way. *Physical Review Letters*, **12**, 155-156. https://doi.org/10.1103/PhysRevLett.12.155

[7] Fung, P.C.W. and Wong, K.W. (2017) Origin of Magnetic Fields of Stellar Objects in the Universe Based on the 5D Projection Theory. *Journal of Modern Physics*, **8**, 668-746. https://doi.org/10.4236/jmp.2017.84045

[8] Fung, P.C.W. and Wong, K.W. (2015) On the Origin of Mass and Angular Momentum of Stellar Objects in the Universe. *Journal of Modern Physics*, **6**, 2303-2341. https://doi.org/10.4236/jmp.2015.615235

[9] Wong, K.W. (1964) Application of Nonlocal Field Operators to a System of Hard Sphere Bose Gas. *Journal of Mathematical Physics*, **5**, 637-642. https://doi.org/10.1063/1.1704157

[10] Aczel, A.D. (1997) Fermat's Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem. Penguin, London, 147 p.

[11] Higgs, P.W. (1964) Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, **18**, 305. https://doi.org/10.1103/PhysRevLett.13.508

[12] Perelman, G. (2002) The Entropy Formula for the Ricci Flow and Its Geometric Applications.

[13] Perelman, G. (2003) Ricci Flow with Surgery on Three-Manifolds.

[14] Ching, W.Y., Xu, Y.N., Zhao, G.-L., Wong, K.W. and Zandiehnadem, F. (1987) Electronic Structure and Excitonic-Enhanced Superconducting Mechanism in $YBa_2Cu_3O_{7-\delta}$. *Physical Review Letters*, **59**, 1333-1336. https://doi.org/10.1103/PhysRevLett.59.1333

[15] Wong, K.W. and Ching, W.Y. (2004) A Structural-Based Microscopic Theory on High-Temperature Cuprate Superconductors. *Physica C*, **416**, 47-67. https://doi.org/10.1016/j.physc.2004.09.003

[16] Zhao, G.L., Xu, Y., Ching, W.Y. and Wong, K.W. (1987) Theoretical Calculation of Optical Properties of Y-Ba-Cu-O Superconductor. *Physical Review B*, **36**, 7203-7206. https://doi.org/10.1103/PhysRevB.36.7203

[17] Wong, K.W. and Ching, W.Y. (1988) Off-Diagonal-Long-Range-Ordering and the Excitonic-Enhancement Model for High Tc Ceramic Oxides. *Proceedings of the* 3*rd Asia Pacific Physics Conference*, Vol. 1, 58-72.

[18] Wong, K.W. and Ching, W.Y. (1988) Thermodynamics of Simultaneous-Excitonic Superconductivity Condensate. *Physica C*, **152**, 397-400. https://doi.org/10.1016/0921-4534(88)90043-3

[19] Wong, K.W. and Ching, W.Y. (1989) The Theory of Simultaneous Excitonic-Superconductivity Condensation. *Physica C*, **158**, 1-14. https://doi.org/10.1016/0921-4534(89)90294-3

[20] Wong, K.W. and Ching, W.Y. (1989) Theory of Simultaneous Excitonic-Superconductivity Condensation II: Experimental Evidences and Stoichiometric Interpretations. *Physica C*, **158**, 15-31. https://doi.org/10.1016/0921-4534(89)90295-5

[21] Wong, K.W., Fung, P.C.W., Yeung, Y.H. and Kwok, W.Y. (1992) Analysis of the Isotope Effect of High-Tc Ceramic $YBa_{2-x}La_xCu_3O_7$ Using the Excitonic-Enhancement

Model. *Physical Review B*, **45**, 13017-13024. https://doi.org/10.1103/PhysRevB.45.13017

[22] Bardeen, J., Cooper, L.N. and Schrieffer, J.R. (1957) Theory of Superconductivity. *Physical Review*, **108**, 1175-1204. https://doi.org/10.1103/PhysRev.108.1175

[23] Wong, K.W. and Curatolo, S. (2008) EEM: The Exciton Enhancement Mechanism Theory and Experimental Evidence of Optically Enhanced Tc in High Tc Superconductors. In: Chang, O.A., Ed., *Progress in Superconductivity Research*, Nova Science Publishers, Inc., Hauppauge, 55-78.

[24] Fan, C.-X. (1990) Dispersive-Mode Theory for the High-Temperature rf Superconducting Quantum Interference Device. *Physical Review B*, **41**, 2041-2045. https://doi.org/10.1103/PhysRevB.41.2041

[25] Johnson, C.Y. (2019) Excessive Brain Activity Linked to a Shorter Life. The Washington Post.

[26] Adams, T., *et al.* (2011) Opera Collaboration.

[27] Lee, T.D. and Yang, C.N. (1956) Question of Parity Conservation in Weak Interactions. *Physical Review*, **104**, 254-258. https://doi.org/10.1103/PhysRev.104.254

Journal of Modern Physics, 2019, 10,1566-1595
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

The Mathematical Foundations of Elasticity and Electromagnetism Revisited

J.-F. Pommaret

CERMICS, Ecole des Ponts Paris Tech, Paris, France
Email: jean-francois.pommaret@wanadoo.fr

How to cite this paper: Pommaret, J.-F. (2019) The Mathematical Foundations of Elasticity and Electromagnetism Revisited. *Journal of Modern Physics*, **10**, 1566-1595.
https://doi.org/10.4236/jmp.2019.1013104

Received: October 1, 2019
Accepted: November 18, 2019
Published: November 21, 2019

Abstract

The first purpose of this striking but difficult paper is to revisit the mathematical foundations of Elasticity (EL) and Electromagnetism (EM) by comparing the structure of these two theories and examining with details their known couplings, in particular *piezoelectricity* and *photoelasticity*. Despite the strange Helmholtz and Mach-Lippmann analogies existing between them, no classical technique may provide a common setting. However, unexpected arguments discovered independently by the brothers E. and F. Cosserat in 1909 for EL and by H. Weyl in 1918 for EM are leading to construct a new differential sequence called *Spencer sequence* in the framework of the formal theory of Lie pseudo groups and to introduce it for the *conformal group* of space-time with 15 parameters. Then, all the previous explicit couplings can be deduced abstractly and one must just go to a laboratory in order to know about the coupling constants on which they are depending, like in the Hooke or Minkowski constitutive relations existing *respectively and separately* in EL or EM. We finally provide a new combined experimental and theoretical proof of the fact that any 1-form with value in the second order jets (*elations*) of the conformal group of space-time can be uniquely decomposed into the direct sum of the Ricci tensor and the electromagnetic field. This result questions the mathematical foundations of both General Relativity (GR) and Gauge Theory (GT). In particular, the *Einstein* operator (6 terms) must be thus replaced by the adjoint of the *Ricci* operator (4 terms only) in the study of gravitational waves.

Keywords

Elasticity, Electromagnetism, Mach-Lippman Analogy, Helmholtz Analogy, Piezoelectricity, Photoelasticity, Differential Sequences, Adjoint Operator, Double Duality

1. Introduction

At the beginning of the last century G. Lippmann and H. von Helmholtz, who

knew each other, were both looking for the possibility to interpret thermostatic and electric phenomena by exhibiting a common macroscopic mechanical origin through a kind of variational calculus similar to the one used in analytical mechanics for getting Euler-Lagrange equations. As a byproduct, it is not possible to separate the *Mach-Lippmann analogy* from the *Helmholtz analogy* that we now recall.

In analytical mechanics, if $L(t,q,\dot{q})$ is the *Lagrangian* of a mechanical system, one easily gets the *Hamiltonian* $H=\dot{q}\frac{\partial L}{\partial \dot{q}}-L$ where t is time, q represents a certain number of dependent variables or *generalized position*, allowing to define the position of the various rigid bodies constituting the system (coordinates of center of gravity, relative angles, ...) and $\dot{q}$ is the derivative with respect to time or *generalized speed*. There are two ideas behind such a construction. The first is to introduce the energy as in the movement of a point of mass m with Cartesian coordinates (x,y,z vertical) or (x^1,x^2,x^3 vertical) in the gravitational field $\boldsymbol{g}$ where $L=\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)-mgz$ and thus $H=\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)+mgz$. The second is to take into account the well known Euler-Lagrange equations $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)-\frac{\partial L}{\partial q}=0$ implied by the variational condition $\delta\int L(t,q,\dot{q})\mathrm{d}t=0$ and to obtain therefore:

$$\frac{\mathrm{d}H}{\mathrm{d}t}=\dot{q}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)-\frac{\partial L}{\partial q}\right)-\frac{\partial L}{\partial t}=-\frac{\partial L}{\partial t}$$

that is the conservation of energy along the trajectories whenever L does not contain t explicitly.

Similarly, in thermostatics, if F is the *free energy* of a system at absolute temperature T, we may obtain, *in general*, the *internal energy* U by the formula $U=F-T\frac{\partial F}{\partial T}$. We explain the underlying difficulty in the case of a perfect gas with pressure P, volume V and entropy S for one mole. The *first principle* of thermostatics says that the sum of the exchange of work $\delta W=-P\mathrm{d}V$ and the exchange of heat δQ between the system and its surrounding is a total differential $\mathrm{d}U=\delta W+\delta Q$. Now, the *second principle* of thermostatics says that $\delta Q=T\mathrm{d}S$ or equivalently that $\frac{\delta Q}{T}=\mathrm{d}S$ is a total differential with absolute temperature as integrating factor. Accordingly, we have $\mathrm{d}U=-P\mathrm{d}V+T\mathrm{d}S$, a result giving U as a function of V and S. As *V has a geometric meaning that S does not possess*, engineers use to do a Legendre transformation by introducing $F=U-TS$ in order to have $\mathrm{d}F=-P\mathrm{d}V-S\mathrm{d}T$ where F is now a function of V and T that can be measured. It follows that $S=-\frac{\partial F}{\partial T}$ *in this situation* because $\delta W=-P\mathrm{d}V$ does not contain $\mathrm{d}T$. Of course, *contrary to S, T can be*

measured though it does not seem to have a geometric meaning like V. In general, the 1-form δW depends linearly on the differentials of *all* the state variables ($\mathrm{d}V$ and $\mathrm{d}T$ in our case) and *there is no reason at all to have again* $S=-\frac{\partial F}{\partial T}$. To avoid such a situation, *Helmholtz postulated the possibility for any system to choose "normal" state variables such that* $\mathrm{d}T$ *should not appear in* δW. Therefore, if one could introduce V and T on an equal geometric footing, then $\mathrm{d}F=-P\mathrm{d}V-S\mathrm{d}T$ should already contain, in a built-in manner, not only the first and second principle but also the well defined possibility to recover U from F as before. In the case of continuum mechanics that we shall study later on, V must be replaced by the deformation tensor, as we shall see later on, which is a function of the first order derivatives of the *actual* (Euler) position x at time t with respect to the *initial* (Lagrange) position x_0 at time t_0. Accordingly, the idea of Helmholtz has been to compare the relations $L\to H$ and $F\to U$ and to notice that they should become indeed similar if one could set $L=-F$ *and* $\dot{q}=T$ *for a certain* q. However, despite many attempts [1], nobody knows any variable q such that its derivative with respect to time should be the absolute temperature T of the system considered.

We now present the work done by Lippmann in a modern setting. The basic idea is to compare two kinds of *conceptual experiments*, namely a Carnot cycle for a steam engine working between the absolute temperatures T_1 and T_2 with $T_2>T_1$ on one side, and a cycle of charge and discharge of a spherical condenser (say a soap buble) of radius r, moving in between two plates at constant electric potentials V_1 and V_2 with $V_2>V_1$ on the other side ([2] [3] [4] [5]).

In the first case, let the system receive the heat $Q_2>0$ from the hot source and the heat $Q_1<0$ from the cold source through corresponding isothermal evolutions, while receiving the work $W<0$ from the surroundings in a cycle completed by two adiabatic evolutions.

The vanishing of the cycle integral:

$$\oint(\delta W+\delta Q)=\oint \mathrm{d}U=0$$

coming from the first principle of thermostatics leads to the relation $W+Q_1+Q_2=0$.

Then, the vanishing of the cycle integral coming from the second principle of thermostatics:

$$\oint\frac{\delta Q}{T}=\oint \mathrm{d}S=0$$

leads to the *Clausius formula* and the computation of the *efficiency* ν:

$$\frac{Q_1}{T_1}+\frac{Q_2}{T_2}=0\Rightarrow \nu=\frac{-W}{Q_2}=\frac{Q_1+Q_2}{Q_2}=\frac{T_2-T_1}{T_2}>0$$

Now, in the second case, *things are quite more subtle*. Recalling the formula $q=CV$ relating the charge q to the potential V of a condenser with $C=4\pi\epsilon_0 r$ for a sphere of radius r, the electric energy should be:

$$E=\frac{1}{2}CV^2=\frac{1}{2}\frac{q^2}{C}=\frac{1}{2}qV$$

Whenever C remains constant, the exchange of work done by the sources should be $\delta W'=V\mathrm{d}q$ because, *by definition*, *sources are at constant potential*, and we have $\mathrm{d}E=q\mathrm{d}V=V\mathrm{d}q=\delta W'$. However, the situation is *completely different* whenever C depends on r and we do not believe that Lippmann was very conscious about this fact. Let us suppose that the bubble receives the work $W_2'>0$ from the source at potential V_2 for having its charge changing at constant potential V_2 and similarly the work $W_1<0$ from the source at constant potential V_1 for having its charge changing at constant potential V_1, while receiving the (mechanical) work $W<0$ from the surroundings for changing C in a cycle where the geometry of the system may vary (change of radius or distance). The problem is now to construct the cycle in order to be able to copy the procedure used for thermostatics. In the evolution at constant potential we have $\delta W'=V\mathrm{d}q$, as already said, and therefore, comparing with $\delta Q=T\mathrm{d}S$, *the remaining evolution must be at constant charge*, a situation happily realized in the experiment proposed by Lippmann, during the transport of the bubble from one plate to the other.

Now, taking into account the expression $\delta W'=V\mathrm{d}q$ already introduced and allowing C to vary (through r in our case), we have the formula:

$$\mathrm{d}E=-\frac{1}{2}V^2\mathrm{d}C+V\mathrm{d}q=\delta W+\delta W'$$

if we express E as a function of C and q. In our case $\delta W=-2\pi\epsilon_0V^2\mathrm{d}r$ and the relation $q=CV$ plays the role of the relation $PV=RT$ existing for a perfect gas.

Copying the use of the first principle of thermostatics, the vanishing of the cycle integral provides:

$$\oint\left(\delta W+\delta W'\right)=\oint\mathrm{d}E=0\Rightarrow W+W_1'+W_2'=0$$

Lippmann then notices that the conservation of entropy now becomes the conservation of charge and the vanishing of the cycle integral provides:

$$\oint\frac{\delta W'}{V}=\oint\mathrm{d}q=0\Rightarrow\frac{W_1'}{V_1}+\frac{W_2'}{V_2}=0\Leftrightarrow\nu=\frac{-W}{W_2'}=\frac{V_2-V_1}{V_2}>0$$

analogous to the Clausius formula with similar *efficiency* ν, a result he called "*Principe de conservation de l'électricité*" or "*Second principe de la théorie des phénomènes électriques*".

One must notice the formula:

$$\mathrm{d}E=\frac{1}{2}V^2\mathrm{d}C+q\mathrm{d}V$$

if we express E as a function of C and V. Also the analogue of the free energy should be $E-qV=-E$ expressed as a function of C and V. Hence *it is not evident, at first sight, to know whether the more* "*geometric*" *quantity is q or V*.

Finally, the analogy between T and V in the corresponding "second principles"

is clear and constitutes the Mach-Lippmann analogy. However, the reader may find strange that T, which is just defined up to a change of scale because of the existence of a reference absolute zero, should be put in correspondence with V which is defined up to an additive constant. In fact, the formula for the spherical condenser (Gauss theorem) is only true if *the potential at infinity is chosen to be zero*, as a zero charge on the sphere is perfectly detectable by counting the number of electrons on the surface. Accordingly, the two previous dimensionless ratios are perfectly well defined, independently of any unit chosen for T or V. However, such an analogy is perfectly coherent with the existence of thermocouples where the gradient of T is proportional to the gradient of V, that is we have for the electric field $\boldsymbol{E}=\eta(T)\boldsymbol{\nabla}T$ and the latter difficulty entirely disappears.

We recall that the *thermoelectric effect*, that is the existence of an electric current circulating in two different metal threads A and B with soldered ends at different temperatures T_1 and $T_2>T_1$, has been discovered in 1821 by the physicist Seebeck from the Netherlands. Also cutting one of the threads to set a condenser and integrating along the circuit, the difference of potential becomes:

$$V=\oint \boldsymbol{E}\cdot \mathrm{d}\ell=\int_{T_1}^{T_2}\left(\eta_A(T)-\eta_B(T)\right)\mathrm{d}T$$

Hence a thermocouple only works if $A\neq B$, $T_1\neq T_2$ and tables of coefficients can be found in the literature. It is the French physicist Becquerel who got the idea in 1830 to use such a property for measuring temperature and Le Chatelier in 1905 who set up the platine thermocouple still used today. Meanwhile, J. Peltier proved that, when an electric current is passing in a thermocouple circuit with soldered joints at the same temperature, then one of the joints absorbs heat while the other produces heat. Also W. Thomson proved that an electric current passing in a piece of homogeneous conductor in thermal equilibrium gives a difference of potential at the ends whenever they are not at the same temperature.

We end this presentation of the Mach-Lippmann analogy with the main problem that it raises. From the special relativity of A. Einstein in 1905 [6] it is known that space cannot be separated from time and that one of the best examples is given by the relativistic formulation of EM. Indeed, instead of writing down separately the first set of Maxwell equations for the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ under their classical form, ne may introduce local coordinates $\left(x^1,x^2,x^3,x^4=ct\right)$ where c is the speed of light and consider the 2-form $F=\left(F_{ij}\right)\in\wedge^2T^*$ with standard notations:

$$F=B_1dx^2\wedge dx^3+B_2dx^3\wedge dx^1+B_3dx^1\wedge dx^2+\frac{1}{c}E_1dx^1\wedge dx^4$$
$$+\frac{1}{c}E_2dx^2\wedge dx^4+\frac{1}{c}E_3dx^3\wedge dx^4$$

in order to obtain:

$$\boldsymbol{\nabla}\cdot\boldsymbol{B}=0,\boldsymbol{\nabla}\wedge\boldsymbol{E}+\frac{\partial\boldsymbol{B}}{\partial t}=0\Leftrightarrow dF=0\Leftrightarrow\partial_iF_{jk}+\partial_jF_{ki}+\partial_kF_{ij}=0$$

where $d:\wedge^2 T^* \to \wedge^3 T^*$ is the exterior derivative.

Similarly, introducing the electromagnetic potential $\boldsymbol{A}$ and the electric potential V in the 1-form $A = A_i dx^i = A_1 dx^1 + A_2 dx^2 + A_3 dx^3 + A_4 dx^4$ where $A_4 = -V/c$ is the time component, we obtain:

$$\boldsymbol{B} = \nabla \wedge \boldsymbol{A}, \quad \boldsymbol{E} = -\nabla V - \frac{\partial \boldsymbol{A}}{\partial t} \Leftrightarrow dA = F$$

though, surprisingly, V has been introduced in thermostatics. Hence, even if we may accept and understand an analogy between T and V, we cannot separate V from $\boldsymbol{A}$ in the 4-potential A and a good conceptual analogy should be between T and $A = (A_1, A_2, A_3, A_4)$.

The surprising fact is that almost nobody knows about the Mach-Lippmann analogy today but many persons are using it through finite element computations and thus any engineer working with finite elements knows that *elasticity*, *heat* and *electromagnetism*, though being quite different theories at first sight, are organized along the same scheme and cannot be separated because of the existence of the following couplings that we shall study with more details in the next Section.

- *THERMOELASTICITY* (Elasticity/Heat):

When a bar of metal is heated, its length is increasing and, conversely, its length is decreasing when it is cooled down. It is a perfectly *reversible phenomenon*.

- *PIEZOELECTRICITY*, *PHOTOELASTICITY* (Elasticity/Electromagnetism):

When a crystal is pinched between the two plates of a condenser, it produces a difference of potential between the plates and conversely, in a purely reversible way. Piezoelectric lighters are of common use in industry. Similarly, when a transparent homogeneous isotropic dielectric is deformed, piezoelectricity cannot appear but the index of refraction becomes different along the three orthogonal proper directions common to both the strain and stress tensors. Here we recall that a material is called "*homogeneous*" if a property does not depend on the point in the material and it is called "*isotropic*" if a property does not depend on the direction in the material. Accordingly, a light ray propagating along one of these directions may have its electric field decomposed along the two others and the two components propagate with different speeds. Hence, after crossing the material, they recompose with production of an interference pattern, a fact leading to optical birefringence. Such a property has been used in order to get information on the stress inside the material, say a bridge or a building, by using reduced transparent plastic models. This phenomenon was discovered by Brewster in 1815 but the phenomenological law that we shall prove in the next section, was proposed independently by F.E. Neumann and J.C. Maxwell in 1830. Until recently one used to rely on the mathematical formulation proposed by Pöckels in 1889 but modern versions can easily be found today in the engineering literature.

- *THERMOELECTRICITY* (Heat/Electromagnetism):

We have already spoken about this coupling which, nevertheless, can only be understood today within the framework of the phenomenological *Onsager relations* for *irreversible phenomena*. If we want to make the Fourier law $\boldsymbol{q}=\chi\nabla T$ between heat flux and gradient of temperature more precise, we may suppose that the heat conductivity χ also depends on the magnetic field and we may obtain "*a priori*" the additional term $q^i=C^{ijkl}F_{kl}\partial_j T$. In a homogeneous medium, one has $C^{ijkl}=a\omega^{ij}\omega^{kl}+b\omega^{ik}\omega^{jl}+c\omega^{il}\omega^{jk}$ when ω is the space Euclidean metric and we have $q^1=(b-c)B_3\partial_2 T$, $q^2=-(b-c)B_3\partial_1 T$ with $\boldsymbol{B}=(0,0,B_3)$, that is $\boldsymbol{q}=(c-b)\boldsymbol{B}\wedge\nabla T$. We have thus been able to recover the Righi-Leduc effect in a purely macroscopic way.

Hence, as a very restrictive conclusion, we discover that the Mach-Lippmann analogy must be at least set up in a clear picture of the analogy existing between elasticity, heat and electromagnetism that must also be coherent with the above couplings.

2. Elasticity versus Electromagnetism

The rough idea is to make the constitutive law of an homogeneous isotropic dielectric $\boldsymbol{D}=\epsilon\boldsymbol{E}$ where $\boldsymbol{D}$ is the electric induction and $\epsilon=\epsilon_0(1+\chi)$, ϵ_0 being the vacuum value (universal constant) of the dielectric constant, such that *the dielectric susceptibility* χ *now depends on the deformation* (*or stress*) *tensor in each direction*. Keeping the constitutive relation $\boldsymbol{H}=\frac{1}{\mu}\boldsymbol{B}$ where $\boldsymbol{H}$ is the magnetic induction and $\mu=\mu_0$ the vacuum value (universal constant) of the magnetic constant, as we have no magnetic polarization in the medium, it is well known that $\epsilon_0\mu_0c^2=1$ and thus $\epsilon\mu c^2=n^2$ where n is the index of refraction such that $n^2=(1+\chi)$, a result leading to the Maxwell-Neumann formula $\sigma_1-\sigma_2=k\lambda/eC$ that we shall demonstrate and apply to the study of a specific beam. In this formula σ_1,σ_2 are the two eigenvalues of the symmetric stress tensor along directions orthogonal to the ray, k is a relative integer fixing the lines of interference, λ is the wave length, e is the thickness of the transparent beam and C is the photoelastic constant of the material.

With more details, the infinitesimal deformation tensor of elasticity theory is equal to half of the Lie derivative $\Omega=(\Omega_{ij}=\Omega_{ji})=\mathcal{L}(\xi)\omega$ of the euclidean metric ω with respect to the displacement vector ξ. Hence, a general quadratic lagrangian may contain, apart from its standard purely elastic or electrical parts well known by engineers in finite element computations, a coupling part $c^{ijk}\Omega_{ij}E_k$ where $E=(E_k)$ is the electric field. The corresponding induction $D=(D^k)$ becomes:

$$D_0^k=\epsilon E^k\rightarrow D^k=D_0^k+c^{ijk}\Omega_{ij}$$

and is therefore modified by an electric polarization $P^k=c^{ijk}\Omega_{ij}$, brought by the deformation of the medium. In all these formulas and in the forthcoming ones the indices are raised or lowered by means of the euclidean metric. If this medium is homogeneous, the components of the 3-tensor c are constants and the corresponding coupling, called *piezoelectricity*, is only existing if the medium is

non-isoptropic (like a crystal), because an isotropic 3-tensor vanishes identically.

In the case of an homogeneous isotropic medium (like a transparent plastic), one must push the coupling part to become cubic by adding $\frac{1}{2}d^{ijkl}\Omega_{ij}E_kE_l$ with $d^{ijkl}=\alpha\omega^{ij}\omega^{kl}+\beta\omega^{ik}\omega^{jl}+\gamma\omega^{il}\omega^{jk}$ from Curie's law. The corresponding coupling, called *photoelasticity*, has been discovered by T. J. Seebeck in 1813 and D. Brewster in 1815. With $\delta=\beta+\gamma$, the new electric induction is:

$$D_0^k=\epsilon E^k\to D^k=D_0^k+\left(\alpha tr(\Omega)\omega^{kr}+\delta\omega^{ik}\omega^{jr}\Omega_{ij}\right)E_r$$

As Ω is a symmetric tensor, we may choose an orthogonal frame at each point of the medium in such a way that the deformation tensor becomes diagonal with $\Omega=(\Omega_1,\Omega_2,\Omega_3)$ where the third direction is orthogonal to the elastic plate. We get:

$$D^i=D_0^i+\left(\alpha tr(\Omega)+\delta\Omega_i\right)E^i$$

for $i=1,2$ without implicit summation and there is a change of the dielectric constant $\epsilon\to\epsilon+\alpha tr(\Omega)+\delta\Omega_i$ along each proper direction in the medium, corresponding to a change $n\to n_i$ of the refraction index. As there is no magnetic property of the medium and $\Omega\ll 1$, we obtain in first approximation:

$$\epsilon\mu_0c^2=n^2\Rightarrow n_1^2-n_2^2\simeq 2n(n_1-n_2)=\mu_0c^2\delta(\Omega_1-\Omega_2)\Rightarrow n_1-n_2\sim\Omega_1-\Omega_2$$

where μ_0 is the magnetic constant of vacuum, c is the speed of light in vacuum and n is the refraction index. The speed of light in the medium becomes c/n_i and therefore depends on the polarization of the beam. As the light is crossing the plate of thickness e put between two polarized filters at right angle, the entering monochromatic beam of light may be decomposed along the two proper directions into two separate beams recovering together after crossing with a time delay equal to:

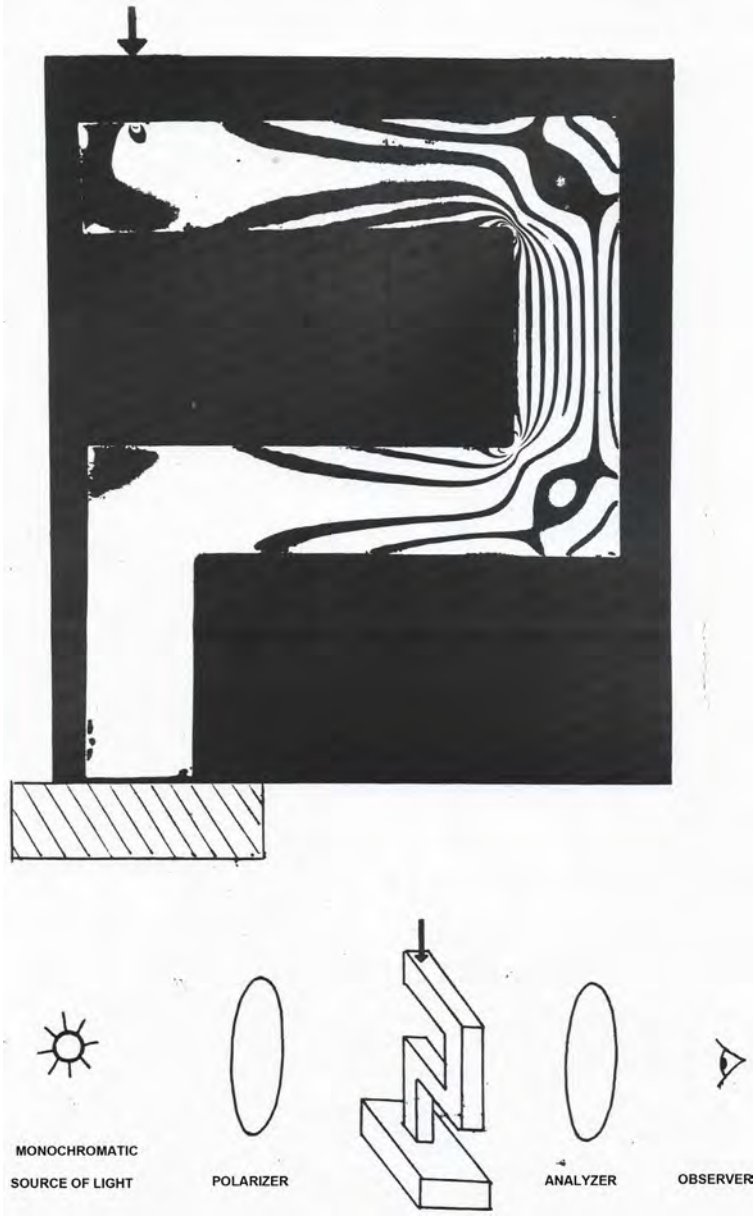
$$e/(c/n_1)-e/(c/n_2)=(e/c)(n_1-n_2)$$

providing interferences and we find back the Maxwell phenomenological law of 1850:

$$\Omega_1-\Omega_2\sim\sigma_1-\sigma_2=\frac{k\lambda}{eC}$$

where σ is the stress tensor, k is an integer, λ is the wave length of the light used and C is the photoelastic constant of the medium involved in the experience.

Looking at the picture, let F be the vertical downwards force acting on the upper left side of the beam like on the picture, at a distance D from the center of the vertical beam on the right. We may consider this vertical beam as a dense sheaf of juxtaposed thin beams with young modulus E. Choosing orthogonal axes ($Oxyz$) such that Ox is horizontal towards the right with origin O in the geometric center of the vertical beam on the right which has a thickness $e=2a$ and a width of $2b$ with the vertical axis Oy passing in the center of the beam. If F should be applied along Oy, according to Hooke's law there should be a vertical compression of the beam providing a deformation roughly equal to



$\epsilon' = -F/(4abE)$ and a (negative because compression) stress $\sigma' = E\epsilon' = -F/4ab$. However, F is applied at a distance D of the axis Oy and gives a couple $M = FD$ which should be, by itself, bringing the half right part of the beam ($x \geq 0$) in extension while the half left part ($x \leq 0$) is in compression. Using a classical assumption usually done on beams we may suppose that the horizontal plane sections orthogonal to the central axis Oy of the beam stay plane surfaces turning counterclockwise by a small angle θ that we shall determine by integration on all the small thin beams of the bunch. The stress σ'' acting on the surface $\mathrm{d}S = e\mathrm{d}x = 2a\mathrm{d}x$ is producing a small force $\mathrm{d}F = \sigma''\mathrm{d}S = 2a\sigma''\mathrm{d}x$. However, a fiber at distance x from the axis has a length increased by θx and there is a resulting deformation $\epsilon'' = Kx$ such that $\sigma'' = E\epsilon''$ on each thin constitutive fiber. The resulting (direct sense)

couple produced is equal to $\mathrm{d}M = x(2aKEx\mathrm{d}x)$ in such a way that we have the equilibrium equation for couples:

$$M = FD = \int_{-b}^{+b}(2aKEx^2)\mathrm{d}x = \frac{4}{3}ab^3KE \quad \Rightarrow \quad K = (3FD)/(4ab^3E)$$

We obtain therefore $\sigma'' = EKx = (3FDx)/(4ab^3)$ with $\sigma'' \geq 0$ whenever $x \geq 0$ (extension). Using the correct negative sign for the stress σ', we finally obtain $\sigma = \sigma^{yy} = \sigma' + \sigma'' = (3FDx)/(4ab^3) - F/(4ab)$ in such a way that $\sigma \leq 0$ when $D = 0$ and $\sigma = 0$ when $x = b^2/3D > 0$, a result not evident at first sight. In addition, it is clear by symmetry that x, y, z are proper directions and that $\sigma^{xx} = \sigma^{zz} = 0$ because no force is acting on the faces of the beam. We obtain therefore the very simple Maxwell law $\sigma = k\lambda/eC$. Accordingly, the (almost!) central black line corresponds to $\sigma = 0$ and has abcissa $x = b^2/3D > 0$. Finally, the distance *d* between two lines is such that *k* is modified by 1, that is $d = (2\lambda b^3)/(3FDC)$, allows to determine the photoelastic constant of the material.

The study of the upper horizontal part of the beam is more delicate. With axis *Oy* in the middle section, starting under the force *F* and axix *Ox* upward, we have $\sigma^{yy} = (3Fxy)/(2eb^3)$ to compensate the couple *Fy* but *now* we have a shear stress $\sigma^{xy} = F/(2eb)$ upward to compensate *F* which is downward. The characteristic polynomial is

$det(\sigma - \lambda\omega) = \lambda^2 - \sigma^{yy}\lambda - (\sigma^{xy})^2 = (\lambda - \sigma^1)(\lambda - \sigma^2) = 0$ and thus $(\sigma^1 - \sigma^2)^2 = (\sigma^1 + \sigma^2)^2 - 4\sigma^1\sigma^2 = (\sigma^{yy})^2 + 4(\sigma^{xy})^2 > 0$ cannot vanish. Therefore the line "$k = 0$" cannot exist. As for the lines "$k = \pm 1$", we must have after substitution $((3Fxy)/2eb^3)^2 + (F/eb)^2 = (\lambda/eC)^2$ and we need to have thus $F < (\lambda b)/C$ or equivalently $d > 2b^2/3D$, a result simply leading to the hyperbola $xy = cst$, a property that can be checked on the picture but cannot be even imagined.

We have thus explained, in a perfectly coherent way with the picture, why the interference lines are parallel and equidistant from each other in the right vertical part of the beam, on both sides of an (*almost*) central line which, *surprisingly*, stops at the upper and lower corner, even though, by continuity, we could imagine that it could be followed in the upper and lower horizontal parts of the beam. Also, we understand now the reason for which the lines in these parts of the beam look like symmetric hyperbolas.

This result proves, *without any doubt for anybody doing this experiment*, that the deformation $\Omega = \mathcal{L}(\xi)\omega$ and the electromagnetic field $F = dA$, using standard notations in the space-time formulation of electromagnetism, *must be on equal footing* in a lagrangian formalism. However, as $\Omega \in S_2T^*$ is in the *Janet sequence* based on the work of E. Vessiot in 1903 (Compare [7] to [8] [9] [10] [11]) and $F \in \wedge^2 T^*$ *cannot* appear at this level as we shall see, the main purpose of this paper is to prove that *another differential sequence must be used*, namely the *Spencer sequence*. The idea has been found *totally independently*, by the brothers E. and F. Cosserat in 1909 ([12]) for revisting elasticity theory and

by H. Weyl in 1916 ([13]) for revisiting electromagnetism by using the conformal group of space time, but the first ones were only dealing with the *translations* and *rotations* while the second was only dealing with the *dilatation* and the non-linear *elations* of this group, with no real progress during the last hundred years.

Extending the space (x^1,x^2,x^3) or (x,y,z) to space-time $(x^1,x^2,x^3,x^4=ct)$ as before, the speed is now extended from (v^1,v^2,v^3) to (v^1,v^2,v^3,c) along the derivative with respect to time, with $v/c \ll 1$, while the motion $x=x_0+\xi(x_0,t)$ is extended to $t=t_0+cst$ in order to compare "slices" of space at the same "time". Accordingly, the deformation tensor ϵ, which is dimensionless, is extended by $(\epsilon_{i4}=\epsilon_{4i})=\frac{1}{2}\left(\frac{v^1}{c},\frac{v^2}{c},\frac{v^3}{c},0\right)$ while the symmetric stress tensor $\sigma^{ij}=\sigma^{ji}$ becomes $\sigma^{ij}-\rho v^i v^j$ (Euler theorem) and is extended by setting $\sigma_{i4}=-\sigma^{i4}=\rho v^i c$, $\sigma^{44}=\sigma_{44}=-\rho c^2$ where ρ is the mass per unit volume. Dealing with the rest-frame and using the (small) dilatation relation $\rho=\rho_0(1-tr\epsilon)$ in which ρ_0 is the value of ρ in the initial position where the body is supposed to be homogeneous, isotropic and unstressed, that is, ρ_0 is supposed to be a constant. The Hooke law is now extended by setting:

$$\sigma_{i4}=2\rho_0 c^2\epsilon_{i4},\ \sigma_{44}+\rho_0 c^2=\rho_0 c^2 tr\epsilon$$

in a way compatible with the conservation of mass and we suddenly discover that *there is no conceptual difference between the Lamé constants* (α,β) (do not confuse the notations) of elasticity and *the magnetic constant* μ *on one side* (*space*) or *the mass per unit volume* ρ *and the dielectric constant* ϵ (time) *on the other side*, all these coupling constants being measured in the reference state in which the body (like vacuum) is homogeneous and isotropic (the index "zero" is omitted for simplicity). This result is perfectly coherent "*a posteriori*" with the analogy existing between the well known formulas for *the speed* v_T *of transverse elastic waves*, *the speed* v_L *of longitudinal elastic waves* or *the speed v of light waves* propagating in a homogeneous isotropic medium, as we have indeed ([14]):

$$v_T=\sqrt{\frac{\beta}{\rho}},\quad v_L=\sqrt{\frac{\alpha+2\beta}{\rho}},\quad v=\sqrt{\frac{1}{\epsilon\mu}}=\sqrt{\frac{1/\mu}{\epsilon}}=\frac{c}{n}$$

We now understand that couplings are in fact more general constitutive laws taking into account the tensorial nature of the various terms involved through the *Curie principle*.

3. General Relativity versus Gauge Theory

Let A be a *unitary ring*, that is $1,a,b\in A\Rightarrow a+b,ab\in A,1a=a1=a$ and even an *integral domain* ($ab=0\Rightarrow a=0$ or $b=0$) with *field of fractions* $K=Q(A)$. However, we shall not always assume that A is commutative, that is *ab* may be different from *ba* in general for $a,b\in A$. We say that $M={}_AM$ is a *left module* over A if $x,y\in M\Rightarrow ax,x+y\in M,\forall a\in A$ or a *right module* M_B

over B if the operation of B on M is $(x,b)\to xb, \forall b\in B$. If M is a left module over A and a right module over B with $(ax)b=a(xb), \forall a\in A, \forall b\in B, \forall x\in M$, then we shall say that $M={}_AM_B$ is a *bimodule*. Of course, $A={}_AA_A$ is a bimodule over itself. We define the *torsion submodule* $t(M)=\{x\in M \mid \exists 0\neq a\in A, ax=0\}\subseteq M$ and M is a *torsion module* if $t(M)=M$ or a *torsion-free module* if $t(M)=0$. We denote by $hom_A(M,A)$ the set of morphisms $f: M\to N$ such that $f(ax)=af(x)$ and set $M^*=hom_A(M,A)$. We shall only consider finitely generated modules, recalling that a sequence of modules and maps is exact if the kernel of any map is equal to the image of the map preceding it ([15]-[20] are good references for homological algebra).

When A is commutative, $hom(M,N)$ is again an A-module for the law $(bf)(x)=f(bx)$ as we have $(bf)(ax)=f(bax)=f(abx)=af(bx)=a(bf)(x)$. In the non-commutative case, things are more complicate and, given ${}_AM$ and ${}_AN_B$, then $hom_A(M,A)$ becomes a right module over B for the law $(fb)(x)=f(x)b$.

DEFINITION 3.1: A module F is said to be *free* if it is isomorphic to a (finite) power of A called the *rank* of F over A and denoted by $rk_A(F)$ while the rank $rk_A(M)$ of a module M is the rank of a maximum free submodule $F\subset M$. It follows from this definition that M/F is a torsion module. In the sequel we shall only consider *finitely presented* modules, namely *finitely generated* modules defined by exact sequences of the type $F_1\xrightarrow{d_1}F_0\xrightarrow{p}M\to 0$ where F_0 and F_1 are free modules of finite ranks m_0 and m_1 often denoted by m and p in examples. A module P is called *projective* if there exists a free module F and another (projective) module Q such that $P\oplus Q\simeq F$.

PROPOSITION 3.2: For any short exact sequence $0\to M'\xrightarrow{f}M\xrightarrow{g}M''\to 0$, we have the important relation $rk_A(M)=rk_A(M')+rk_A(M'')$, even in the non-commutative case. As a byproduct, if M admits a finite length free *resolution* $\cdots\xrightarrow{d_2}F_1\xrightarrow{d_1}F_0\xrightarrow{p}M\to 0$, we may define the *Euler-Poincaré characteristic* $\chi_A(M)=\sum_r(-1)^r rk_A(F_r)=rk_A(M)$.

We now turn to the operator framework with modules over the ring $D=K[d_1,\cdots,d_n]=K[d]$ of differential operators with coefficients in a differential field K with n commuting derivations $(\partial_1,\cdots,\partial_n)$, also called D-modules. Then D is a differential bimodule over itself ([18] [19] [21] [22] [23] [24] are good references for differential homological algebra while [8] [9] [10] [11] [18] [19] [25] are good references for the formal theory of systems of partial differential equations).

DEFINITION 3.3: If a differential operator $\xi\xrightarrow{\mathcal{D}}\eta$ is given, a *direct problem* is to find generating *compatibility conditions* (CC) as an operator $\eta\xrightarrow{\mathcal{D}_1}\zeta$ such that $\mathcal{D}\xi=\eta\Rightarrow\mathcal{D}_1\eta=0$. Conversely, given $\eta\xrightarrow{\mathcal{D}_1}\zeta$, the *inverse problem* will be to look for $\xi\xrightarrow{\mathcal{D}}\eta$ such that $\mathcal{D}_1$ generates the CC of $\mathcal{D}$ and we shall say that $\mathcal{D}_1$ *is parametrized by* $\mathcal{D}$ *if such an operator* $\mathcal{D}$ *is existing*.

Introducing the morphism $\epsilon : M \to M^{**}$ such that $\epsilon(m)(f) = f(m), \forall m \in M, \forall f \in M^*$ and defining the differential module N from $ad(\mathcal{D}_1)$ exactly like we defined the differential module M from $\mathcal{D}$, we finally notice that any operator is the adjoint of a certain operator because $ad(ad(P)) = P, \forall P \in D$ and we get ([18] [19] [26] [27] [28] [29]):

THEOREM 3.4: (*double differential duality test*) In order to check whether M is *torsion-free* or not, that is to find out a parametrization if $t(M) = 0$, the test has 5 steps which are drawn in the following diagram where $ad(\mathcal{D})$ generates the CC of $ad(\mathcal{D}_1)$ and $\mathcal{D}_1'$ generates the CC of $\mathcal{D} = ad(ad(\mathcal{D}))$:

$$\begin{array}{ccccccccc} & & & & & & & \zeta' & 5 \\ & & & & & & \overset{\mathcal{D}_1'}{\nearrow} & & \\ 4 & \xi & \overset{\mathcal{D}}{\to} & \eta & \overset{\mathcal{D}_1}{\to} & \zeta & & 1 \\ & & & & & & & & \\ 3 & \nu & \overset{ad(\mathcal{D})}{\leftarrow} & \mu & \overset{ad(\mathcal{D}_1)}{\leftarrow} & \lambda & & 2 \end{array}$$

$$\mathcal{D}_1 \text{ parametrized by } \mathcal{D} \Leftrightarrow \mathcal{D}_1 \simeq \mathcal{D}_1' \Leftrightarrow \epsilon \text{ injective} \Leftrightarrow t(M) = 0$$

COROLLARY 3.5: In the differential module framework, if $F_1 \xrightarrow{\mathcal{D}_1} F_0 \xrightarrow{p} M \to 0$ is a finite free presentation of $M = coker(\mathcal{D}_1)$ with $t(M) = 0$, then we may obtain an exact sequence $F_1 \xrightarrow{\mathcal{D}_1} F_0 \xrightarrow{\mathcal{D}} E$ of free differential modules where $\mathcal{D}$ is the parametrizing operator. However, there may exist other parametrizations $F_1 \xrightarrow{\mathcal{D}_1} F_0 \xrightarrow{\mathcal{D}'} E'$ called *minimal parametrizations* such that $coker(\mathcal{D}')$ is a torsion module and we have thus $rk_D(M) = rk_D(E')$.

These results have been used in *control theory* and it is now known that a control system is *controllable* if and only if it is parametrizable (See [18] [19] [30] [31] for more details). As a byproduct, and *though it is still not acknowledged by engineers*, controllability is a "*built in*" property that does not depend *at all* on the choice of the inputs and outputs among the control variables.

Keeping the same "*operational*" notations for simplicity, we may state ([18], p 638-650):

DEFINITION 3.6: We say that $\mathcal{D} : \xi \to \eta$ admits a (*generalized*) *lift* $P : \eta \to \xi$ if $\mathcal{D} \circ \mathcal{P} \circ \mathcal{D} = \mathcal{D}$. The differential module determined by $\mathcal{D}$ is projective if and only if $\mathcal{D}$ admits a lift.

The following results have never been used for applications:

LEMMA 3.7: If $\mathcal{D}$ admits a lift, then $ad(\mathcal{D})$ also admits a lift.

PROPOSITION 3.8: If $\mathcal{D}$ parametrizes $\mathcal{D}_1$ and admits a lift $\mathcal{P}$, then $\mathcal{D}_1$ admits a lift $\mathcal{P}_1$ and we have the striking *Bezout identity* $\mathcal{D} \circ \mathcal{P} + \mathcal{P}_1 \circ \mathcal{D}_1 = id_\eta$. Accordingly, the corresponding differential sequence, which is formally exact by definition, is also locally exact.

COROLLARY 3.9: If $\mathcal{D}_1$ generates the CC of $\mathcal{D}$ and both operators admit lifts, then $ad(\mathcal{D})$ generates the CC of $ad(\mathcal{D}_1)$.

EXAMPLE 3.10: With $n = 2$, $m = 2$, $q = 1$, $a \in K = \mathbb{Q}(x^1, x^2)$,

$D=K[d_1,d_2]$, $D\eta=D\eta^1+D\eta^2$ and $\Phi\equiv d_1\eta^1+d_2\eta^2-a\eta^1$ we shall prove that $M_1=D\eta/D\Phi$ is torsion-free but not projective when $\partial_2 a=0$ and projective but not free when $\partial_2 a\neq 0$, for example when $a=x^2$. Multiplying Φ by a test function λ and integrating by parts formally the equation $\mathcal{D}_1\eta=\zeta$, we get the operator $ad(\mathcal{D}_1)$ in the form:

$$-d_1\lambda-a\lambda=\mu^1,\ -d_2\lambda=\mu^2\Rightarrow(\partial_2 a)\lambda=d_1\mu^2-d_2\mu^1+a\mu^2$$

- $\partial_2 a=0$: We get the only generating CC $d_1\mu^2-d_2\mu^1+a\mu^2=0$ and $ad(\mathcal{D}_1)$ is not injective. There is therefore no lift and thus no splitting. Multilying by a test function ϕ and integrating by parts, we obtain the parametrization $\mathcal{D}:\phi\to\xi$ in the form $d_2\phi=y^1,-d_1\phi+a\phi=y^2$ which is not injective. The corresponding sequence $D\xrightarrow{\mathcal{D}_1}D^2\xrightarrow{\mathcal{D}}D$ with differential modules and its formal adjoint are both formally exact.
- $\partial_2 a\neq 0$: The situation is now *totally different*. In order to prove this, if we suppose that $a=x^2$, we get the lift $\lambda=d_1\mu^2-d_2\mu^1+x^2\mu^2$ with adjoint $d_2\zeta=\eta^1,-d_1\zeta+x^2\zeta=\eta^2$ providing a lift $\mathcal{P}_1$ for $\mathcal{D}_1$. Substituting, we obtain two second-order CC ν^1 and ν^2 satisfying the only CC $d_1\nu^2-d_2\nu^1+x^2\nu^2=0$. Multiplying these two CC by the test functions ξ^1 and ξ^2 and integrating by parts, we finally obtain the involutive parametrizing operator $\mathcal{D}$ in the form:

$$d_{12}\xi^1+d_{22}\xi^2-x^2d_2\xi^1-2\xi^1=\eta^1,$$
$$-d_{11}\xi^1-d_{12}\xi^2+2x^2d_1\xi^1+x^2d_2\xi^2-(x^2)^2\xi^1-\xi^2=\eta^2$$

and "*a*" minimum involutive parametrization (but there can be others!):

$$d_{22}\xi=\eta^1,\ -d_{12}\xi+x^2d_2\xi-\xi=\eta^2$$

We get the long formally and locally exact differential sequence $0\to D\xrightarrow{\mathcal{D}_{-1}}D^2\xrightarrow{\mathcal{D}}D^2\xrightarrow{\mathcal{D}_1}D\to 0$ and invite the reader to find a lift for the central operator as an exercise.

EXAMPLE 3.11: When $n=3$, the *div* operator can be parametrized by the *curl* operator which can be itself parametrized by the *grad* operator. However, using $(\xi^1,\xi^2,\xi^3=0)$, we may obtain the new minimal parametrization $-\partial_3\xi^2=\eta^1$, $\partial_3\xi^1=\eta^2$, $\partial_1\xi^2-\partial_2\xi^1=\eta^3\Rightarrow\partial_1\eta^1+\partial_2\eta^2+\partial_3\eta^3=0$ which cannot be again parametrized ([29] [32]).

EXAMPLE 3.12: *Parametrization of the Cauchy stress equations.*

We shall consider the cases $n=2,3,4$ but the case *n* arbitrary could be treated as well.

- $n=2$: The stress equations become $\partial_1\sigma^{11}+\partial_2\sigma^{12}=0$, $\partial_1\sigma^{21}+\partial_2\sigma^{22}=0$. Their second order parametrization $\sigma^{11}=\partial_{22}\phi$, $\sigma^{12}=\sigma^{21}=-\partial_{12}\phi$, $\sigma^{22}=\partial_{11}\phi$ has been provided by George Biddell Airy in 1863 ([33]) and we shall thus denote by $Airy:\phi\to\sigma$ the corresponding operator. We get the linear second order system with formal notations:

$$\begin{cases}\sigma^{11}\equiv d_{22}\phi=0\\ -\sigma^{12}\equiv d_{12}\phi=0\\ \sigma^{22}\equiv d_{11}\phi=0\end{cases}\quad\begin{array}{|cc|}\hline 1&2\\ 1&\bullet\\ 1&\bullet\\ \hline\end{array}$$

which is involutive with one equation of class 2, 2 equations of class 1 and it is easy to check that the 2 corresponding first order CC is just the stress equations. Now, multiplying the Cauchy stress equations respectively by test functions ξ^1 and ξ^2, then integrating by parts, we discover that (up to sign and a factor 2) the *Cauchy operator* is the formal adjoint of the *Killing operator* defined by $\mathcal{D}\xi = \mathcal{L}(\xi)\omega = \Omega \in S_2T^*$, introducing the standard Lie derivative of the (non-degenerate) euclidean metric ω with respect to ξ and using the fact that we have $\sigma^{ij}\Omega_{ij} = \sigma^{11}\Omega_{11} + 2\sigma^{12}\Omega_{12} + \sigma^{22}\Omega_{22}$ *because we have supposed that* $\sigma^{12} = \sigma^{21}$ and we shall say, with a slight abuse of language, that $Cauchy = ad(Killing)$. In order to apply the above parametrization test, we have to look for the CC $\mathcal{D}_1$ of $\mathcal{D}$. In arbitrary dimension *n*, we may introduce the Riemann tensor $\rho^k_{l,ij}$ with $n^2(n^2-1)/12$ components of a general metric ω such that $det(\omega) \neq 0$ and linearize it over a given non-degenerate constant metric or, more generally, over a metric with constant Riemaniann curvature, in order to obtain the second order *Riemann* operator $(\Omega_{ij}) \to (R^k_{l,ij})$. When $n = 2$ and ω is the euclidean metric, we get a single component that can be chosen to be the scalar curvature $R = d_{11}\Omega_{22} + d_{22}\Omega_{11} - 2d_{12}\Omega_{12}$. Multiplying by a test function ϕ and integrating by parts, we obtain $Airy = ad(Riemann)$ and notice that:

There is no relation at all between the Airy stress function ϕ *and the deformation* Ω *of the metric* ω.

- $n = 3$: Things become quite more delicate when we try to parametrize the 3 PD equations:

$$\begin{aligned}
&\partial_1\sigma^{11} + \partial_2\sigma^{12} + \partial_3\sigma^{13} = 0,\\
&\partial_1\sigma^{21} + \partial_2\sigma^{22} + \partial_3\sigma^{23} = 0,\\
&\partial_1\sigma^{31} + \partial_2\sigma^{32} + \partial_3\sigma^{33} = 0
\end{aligned}$$

A direct computational approach has been provided by Eugenio Beltrami in 1892 ([34]), James Clerk Maxwell in 1870 ([35]) and Giacinto Morera in 1892 ([36]) by introducing 6 *stress functions* $\phi_{ij} = \phi_{ji}$ in the *Beltrami parametrization* described by the following *Beltrami* operator:

$$\begin{pmatrix} \sigma^{11} \\ \sigma^{12} \\ \sigma^{13} \\ \sigma^{22} \\ \sigma^{23} \\ \sigma^{33} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & d_{33} & -2d_{23} & d_{22} \\ 0 & -d_{33} & d_{23} & 0 & d_{13} & -d_{12} \\ 0 & d_{23} & -d_{22} & -d_{13} & d_{12} & 0 \\ d_{33} & 0 & -2d_{13} & 0 & 0 & d_{11} \\ -d_{23} & d_{13} & d_{12} & 0 & -d_{11} & 0 \\ d_{22} & -2d_{12} & 0 & d_{11} & 0 & 0 \end{pmatrix} \begin{pmatrix} \Phi_{11} \\ \Phi_{12} \\ \Phi_{13} \\ \Phi_{22} \\ \Phi_{23} \\ \Phi_{33} \end{pmatrix}$$

It is involutive with 3 equations of class 3, 3 equations of class 2 and no equation of class 1. The 3 CC is describing the stress equations which admit therefore a parametrization, but without any geometric framework, in particular without any possibility to imagine that the above second order operator is *nothing else but* the *formal adjoint* of the *Riemann operator*, namely the (linearized) Riemann tensor with $n^2(n^2-1)/12 = 6$ independent components when $n = 3$ [8]

[9] [10] [11]. We may rewrite the Beltrami parametrization of the Cauchy stress equations as follows, after exchanging the third row with the fourth row and using formal notations:

$$\begin{pmatrix} d_1 & d_2 & d_3 & 0 & 0 & 0 \\ 0 & d_1 & 0 & d_2 & d_3 & 0 \\ 0 & 0 & d_1 & 0 & d_2 & d_3 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & d_{33} & -2d_{23} & d_{22} \\ 0 & -d_{33} & d_{23} & 0 & d_{13} & -d_{12} \\ 0 & d_{23} & -d_{22} & -d_{13} & d_{12} & 0 \\ d_{33} & 0 & -2d_{13} & 0 & 0 & d_{11} \\ -d_{23} & d_{13} & d_{12} & 0 & -d_{11} & 0 \\ d_{22} & -2d_{12} & 0 & d_{11} & 0 & 0 \end{pmatrix} \equiv 0$$

as an identity where 0 on the right denotes the zero operator. However, the standard implicit summation used in continuum mechanics (See [36] for more details) is, when $n=3$:

$$\sigma^{ij}\Omega_{ij} = \sigma^{11}\Omega_{11} + 2\sigma^{12}\Omega_{12} + 2\sigma^{13}\Omega_{13} + \sigma^{22}\Omega_{22} + 2\sigma^{23}\Omega_{23} + \sigma^{33}\Omega_{33}$$

because *the stress tensor density* σ *is supposed to be symmetric* in continuum mechanics. Integrating by parts in order to construct the adjoint operator, we get the striking identification:

$$Riemann = ad\left(Beltrami\right) \quad \Leftrightarrow \quad Beltrami = ad\left(Riemann\right)$$

between the (linearized) Riemann tensor and the Beltrami parametrization.

As we already said, the brothers E. and F. Cosserat proved in 1909 that the assumption $\sigma^{ij} = \sigma^{ji}$ may be too strong because it only takes into account density of forces and ignores density of couples, and the Cauchy stress equations *must* be replaced by the so-called *Cosserat couple-stress equations* ([9] [10] [12] [37] [38]). In any case, taking into account the factor 2 involved by multiplying the second, third and fifth row by 2, we get the new 6×6 matrix with rank 3:

$$\begin{pmatrix} 0 & 0 & 0 & d_{33} & -2d_{23} & d_{22} \\ 0 & -2d_{33} & 2d_{23} & 0 & 2d_{13} & -2d_{12} \\ 0 & 2d_{23} & -2d_{22} & -2d_{13} & 2d_{12} & 0 \\ d_{33} & 0 & -2d_{13} & 0 & 0 & d_{11} \\ -2d_{23} & 2d_{13} & 2d_{12} & 0 & -2d_{11} & 0 \\ d_{22} & -2d_{12} & 0 & d_{11} & 0 & 0 \end{pmatrix}$$

This is a symmetric matrix and *the corresponding second order operator with constant coefficients is thus self-adjoint*.

Surprisingly, the Maxwell parametrization is obtained by keeping only $\phi_{11} = A$, $\phi_{22} = B$, $\phi_{33} = C$ while setting $\phi_{12} = \phi_{23} = \phi_{31} = 0$ and using only the columns $1+4+6$ as follows:

$$\begin{pmatrix} \sigma^{11} \\ \sigma^{12} \\ \sigma^{13} \\ \sigma^{22} \\ \sigma^{23} \\ \sigma^{33} \end{pmatrix} = \begin{pmatrix} 0 & d_{33} & d_{22} \\ 0 & 0 & -d_{12} \\ 0 & -d_{13} & 0 \\ d_{33} & 0 & d_{11} \\ -d_{23} & 0 & 0 \\ \partial_{22} & d_{11} & 0 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix}$$

and we let the reader check the corresponding Cauchy equations.

$n=4$: It is only now that we are able to explain the relation of this striking result with Einstein equations but the reader must already understand that, if we need to revisit in such a deep way the mathematical foundations of elasticity theory, we also need to revisit in a similar way the mathematical foundations of EM and GR as in ([28] [32] [36] [39] [40] [41] [42]). To begin with, let us introduce the *Ricci* operator $(\Omega_{ij})\to(R_{ij})$ with 4 terms and the *Einstein* operator $(\Omega_{ij})\to\left(E_{ij}=R_{ij}-\frac{1}{2}\omega_{ij}tr(R)\right)$ with 6 terms where the trace of (R_{ij}) is just $tr(R)=\omega^{ij}R_{ij}$. Surprisingly, the *Einstein* operator is self adjoint while the *Ricci* operator is not and "*Einstein equations are just a way to parametrize the Cauchy stress equations*" because of the well known contraction of the Bianchi identities ([29] [31] [32] [44]). Now, Theorem 3.4 proves that the *Einstein* operator cannot be parametrized ([31] [40]) and that each component of the Weyl tensor is a torsion element killed by the Dalembertian ([18] [32] [44]). We now prove that only the use of *differential homological algebra*, a mixture of differential geometry (differential sequences, formal adjoint) and homological algebra (module theory, double duality, extension modules) *totally unknown by physicists*, is able to explain why the Einstein operator (with 6 terms) defined above is useless as it can be replaced by the Ricci operator (with 4 terms) in the search for gravitational waves equations. Indeed, denoting by $\Omega\in S_2T^*$ a perturbation of the non-degenerate metric ω, it is well known (See [8] [10] and [47] for more details) that the linearization of the *Ricci* tensor $R=(R_{ij})\in S_2T^*$ over the Minkowski metric, considered as a second order operator $\Omega\to R$, may be written *with four terms* as ([32] [43]):

$$2R_{ij}=\omega^{rs}\left(d_{ij}\Omega_{rs}+d_{rs}\Omega_{ij}-d_{ri}\Omega_{sj}-d_{sj}\Omega_{ri}\right)=2R_{ji}$$

Multiplying by test functions $(\lambda^{ij})\in\wedge^4T^*\otimes S_2T$ and integrating by parts on space-time, we obtain the following *four terms* describing the so-called *gravitational waves equations*:

$$\left(\Box\lambda^{rs}+\omega^{rs}d_{ij}\lambda^{ij}-\omega^{sj}d_{ij}\lambda^{ri}-\omega^{ri}d_{ij}\lambda^{sj}\right)\Omega_{rs}=\sigma^{rs}\Omega_{rs}$$

where $\Box$ is the standard Dalembertian. Accordingly, we have:

$$d_r\sigma^{rs}=\omega^{ij}d_{rij}\lambda^{rs}+\omega^{rs}d_{rij}\lambda^{ij}-\omega^{sj}d_{rij}\lambda^{ri}-\omega^{ri}d_{rij}\lambda^{sj}=0$$

The basic idea used in GR has been to simplify these equations by adding the *differential constraints* $d_r\lambda^{rs}=0$ in order to find only $\Box\lambda^{rs}=\sigma^{rs}$, exactly like in the Lorenz condition for EM. It follows that the *Cauchy* = *ad*(*Killing*) operator is parametrized by $ad(Ricci)$ and, *not only* the Einstein operator is useless as it must be replaced by $ad(Ricci)$ *but also* this result shows that the *Cauchy* operator has nothing to do with the *Bianchi* operator. Finally, as our comment on the *Airy* operator when $n=2$ is still valid, λ *has nothing to do with* Ω and we may say ([32]):

These purely mathematical results question the origin and existence of gravitational waves.

It remains to prove that, in this new framework, the Ricci tensor only depends on the symbol $\hat{g}_2 \simeq T^* \subset S_2T^* \otimes T$ of the first prolongation $\hat{R}_2 \subset J_2(T)$ of the conformal Killing system $\hat{R}_1 \subset J_1(T)$ with symbol $\hat{g}_1 \subset T^* \otimes T$ defined by the equations $\omega_{rj}\xi^r_i + \omega_{ir}\xi^r_j - \frac{2}{n}\omega_{ij}\xi^r_r = 0$ *not depending on any conformal factor*. In the next general commutative diagram covering both situations while taking into account that the PD equations of both the classical and conformal Killing systems are homogeneous, the *Spencer map* δ is induced by minus the Spencer operator and all the sequences are exact but perhaps the left column with δ-cohomology $H^2(g_1) \neq 0$ at $\wedge^2 T^* \otimes g_1$ (See [8] [9] [10] [11] or [25] for more details):

$$
\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 \rightarrow & & g_3 & \rightarrow & S_3T^* \otimes T & \rightarrow & S_2T^* \otimes F_0 & \rightarrow & F_1 \rightarrow 0 \\
 & & \downarrow \delta & & \downarrow \delta & & \downarrow \delta & & \\
0 \rightarrow & T^* \otimes g_2 & \rightarrow & T^* \otimes S_2T^* \otimes T & \rightarrow & T^* \otimes T^* \otimes F_0 & \rightarrow & 0 & \\
 & \downarrow \delta & & \downarrow \delta & & \downarrow \delta & & & \\
0 \rightarrow & \wedge^2 T^* \otimes g_1 & \rightarrow & \underline{\wedge^2 T^* \otimes T^* \otimes T} & \rightarrow & \wedge^2 T^* \otimes F_0 & \rightarrow & 0 & \\
 & \downarrow \delta & & \downarrow \delta & & \downarrow & & & \\
0 \rightarrow & \wedge^3 T^* \otimes T & = & \wedge^3 T^* \otimes T & \rightarrow & 0 & & & \\
 & \downarrow & & \downarrow & & & & & \\
 & 0 & & 0 & & & & &
\end{array}
$$

We have the following fiber dimensions for the classical Killing case and arbitrary dimension n:

$$
\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & 0 & \rightarrow & n^2(n+1)(n+2)/6 & \rightarrow & n^2(n+1)^2/4 & \rightarrow & n^2(n^2-1)/12 \rightarrow 0 \\
 & & & & \downarrow \delta & & \downarrow \delta & & \\
 & & 0 & \rightarrow & n^3(n+1)/2 & \rightarrow & n^3(n+1)/2 & \rightarrow & 0 \\
 & & \downarrow & & \downarrow \delta & & \downarrow \delta & & \\
0 \rightarrow & & n^2(n-1)^2/4 & \rightarrow & n^3(n-1)/2 & \rightarrow & n^2(n^2-1)/4 & \rightarrow & 0 \\
 & & \downarrow \delta & & \downarrow \delta & & \downarrow & & \\
0 \rightarrow & n^2(n-1)(n-2)/6 & = & n^2(n-1)(n-2)/6 & \rightarrow & 0 & & & \\
 & \downarrow & & \downarrow & & & & & \\
 & 0 & & 0 & & & & &
\end{array}
$$

allowing to recover the number of components of the Riemann tensor... without indices ! with $n^2(n-1)^2/4 - n^2(n-1)(n-2)/6 = n^2(n^2-1)/12$ too.

We obtain at once from a snake-type chase the isomorphism $F_1 \simeq H^2(g_1)$ and provide a new simple proof of the following important result (Compare to [10] [11] [32] [40] [41] and the Remark below):

THEOREM 3.13: Introducing the δ-cohomologies $H^2(g_1)$ at $\wedge^2 T^* \otimes g_1$ and $H^2(\hat{g}_1)$ at $\wedge^2 T^* \otimes \hat{g}_1$ while taking into account that $g_1 \subset \hat{g}_1$, we have

the short exact sequences:

$$0\to S_2T^*\to H^2\left(g_1\right)\to H^2\left(\hat{g}_1\right)\to 0 \quad\Leftrightarrow\quad 0\to S_2T^*\to F_1\to \hat{F}_1\to 0$$

Proof: The first result can be deduced from a delicate unusual chase in the following commutative diagram where only the rows and the right column are short exact sequences. The first step is made by a diagonal snake-type chase for defining the left morphism and we let the reader check that it is a monomorphism. The right morphism is described by the inclusion $\wedge^2T^*\otimes g_1\subset\wedge^2T^*\otimes\hat{g}_1$ induced by the inclusion $g_1\subset\hat{g}_1$ by showing that any element of $\wedge^2T^*\otimes\hat{g}_1$ is a sum of an element in $\wedge^2T^*\otimes g_1$ plus the image by δ of an element in $T^*\otimes\hat{g}_2$ for the right epimorphism (exercise).

$$\begin{array}{ccccccccc}
 & & & & & & 0 & & \\
 & & & & & & \downarrow & & \\
 & & & & 0 & & S_2T^* & & \\
 & & & & \downarrow & & \downarrow\delta & & \\
 & & 0 & \to & T^*\otimes\hat{g}_2 & \to & T^*\otimes T^* & \to & 0 \\
 & & \downarrow & & \downarrow\delta & & \downarrow\delta & & \\
0 & \to & \wedge^2T^*\otimes g_1 & \to & \wedge^2T^*\otimes\hat{g}_1 & \to & \wedge^2T^* & \to & 0 \\
 & & \downarrow\delta & & \downarrow\delta & & \downarrow & & \\
0 & \to & \wedge^3T^*\otimes T & = & \wedge^3T^*\otimes T & \to & 0 & & \\
 & & \downarrow & & \downarrow & & & & \\
 & & 0 & & 0 & & & &
\end{array}$$

with fiber dimensions when $n\geq 4$:

$$\begin{array}{ccccccccc}
 & & & & & & 0 & & \\
 & & & & & & \downarrow & & \\
 & & & & 0 & & n(n+1)/2 & & \\
 & & & & \downarrow & & \downarrow\delta & & \\
 & & 0 & \to & n^2 & \to & n^2 & \to & 0 \\
 & & \downarrow & & \downarrow\delta & & \downarrow\delta & & \\
0 & \to & n^2(n-1)^2/4 & \to & n(n-1)(n^2-n+2)/4 & \to & n(n-1)/2 & \to & 0 \\
 & & \downarrow\delta & & \downarrow\delta & & \downarrow & & \\
0 & \to & n^2(n-1)(n-2)/6 & = & n^2(n-1)(n-2)/6 & \to & 0 & & \\
 & & \downarrow & & \downarrow & & & & \\
 & & 0 & & 0 & & & &
\end{array}$$

Using the previous diagram, we obtain the isomorphisms $F_1\simeq H^2\left(g_1\right)$ and $\hat{F}_1\simeq H^2\left(\hat{g}_1\right)$. We have thus the splitting sequence $0\to S_2T^*\to F_1\to\hat{F}_1\to 0$ providing a totally unusual interpretation of the successive Ricci, Riemann and Weyl tensors. It follows that $dim\left(\hat{F}_1\right)=n(n+1)(n+2)(n-3)/12$ whenever $n\geq 4$ and the *Weyl*-type operator is of order 3 when $n=3$ but of order 2 for $n\geq 4$. Similar results could be obtained for the *Bianchi*-type operator ... with much more work!

Q.E.D

REMARK 3.14: Using the contraction $T^*\otimes T\to\wedge^0T^*\to 0$, namely $\xi_i^k\to\xi_r^r$, in order to describe the cokernel of the left vertical monomorphism, we obtain the following commutative and exact diagram which is only depending on the first order jets of T:

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & & & \\
 & & \downarrow & & \downarrow & & & & \\
0 & \to & g_1 & \to & T^*\otimes T & \to & F_0 & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & \hat{g}_1 & \to & T^*\otimes T & \to & \hat{F}_0 & \to & 0 \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

Prolonging twice to the jets of order 3 of T, we obtain the commutative and exact diagram:

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & 0 & \to & S_2T^* & = & S_2T^* & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & S_3T^*\otimes T & \to & S_2T^*\otimes F_0 & \to & F_1 & \to & 0 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \to & S_3T^*\otimes T & \to & S_2T^*\otimes \hat{F}_0 & \to & \hat{F}_1 & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

providing the same short exact sequence as in the Theorem but without any possibility to establish a link between S_2T^* and a 1-form with value in the bundle $\hat{g}_2$ of elations.

EXAMPLE 3.15: *Electromagnetism*.

Passing now to electromagnetism and the original Gauge Theory (GT) which is still, up to now, the only known way to establish a link between EM and group theory, the first idea is to introduce the *nonlinear gauge sequence*:

$$\begin{array}{ccccc}
X\times G & \to & T^*\otimes\mathcal{G} & \overset{MC}{\to} & \wedge^2T^*\otimes\mathcal{G} \\
a & \to & a^{-1}da=A & \to & dA-[A,A]=F
\end{array}$$

where X is a manifold, G is a Lie group with identity e *not acting on* X, $a:X\to G$ a map identified with a section of the trivial bundle $X\times G$ over X and $a^{-1}da=A$ is the pull-back over X by the tangent mapping $T(a)$ of a basis of left invariant 1-forms on G. Also, $[A(\xi),A(\eta)]\in\mathcal{G},\forall\xi,\eta\in T$ by introducing the bracket on the Lie algebra $\mathcal{G}=T_e(G)$ and the pull-back of the Maurer-Cartan (MC) equations on G is the so-called *Cartan curvature* 2-form with value in $\mathcal{G}$. Choosing a close to e, that is $a(x)=e+t\lambda(x)+\cdots$ with $t\ll 1$ and linearizing as usual, we obtain the linear operator $d:\wedge^0T^*\otimes\mathcal{G}\to\wedge^1T^*\otimes\mathcal{G}:(\lambda^\tau(x))\to(\partial_i\lambda^\tau(x))$ leading to the *linear gauge sequence*:

$$\wedge^0 T^* \otimes \mathcal{G} \xrightarrow{d} \wedge^1 T^* \otimes \mathcal{G} \xrightarrow{d} \wedge^2 T^* \otimes \mathcal{G} \xrightarrow{d} \cdots \xrightarrow{d} \wedge^n T^* \otimes \mathcal{G} \to 0$$

which is the tensor product by $\mathcal{G}$ of the Poincaré sequence for the exterior derivative *d*. In 1954, at the birth of GT, the above notations were coming from electromagnetism with EM *potential* $A \in T^*$ and EM *field* $dA = F \in \wedge^2 T^*$ in the relativistic Maxwell theory. Accordingly, $G = U(1)$ (unit circle in the complex plane) $\to dim(\mathcal{G}) = 1$ was the *only possibility* existing before 1970 to get a *pure* 1-form *A* (EM potential) and a *pure* 2-form *F* (EM field) when *G* is abelian. However, this result is *not coherent at all* with elasticity theory as we saw and, *a fortiori*, with the analytical mechanics of rigid bodies where the Lagrangian is a quadratic expression of such 1-forms when $n = 3$ and $G = SO(3)$ (Compare to [45] and [46]).

Before going ahead, let us prove that there may be mainly two types of differential sequences, the *Janet sequence* introduced by M. Janet in 1970 ([8] [11] [47]) for the dealing with successive *compatibility conditions* (CC), and a quite different sequence called *Spencer sequence* introduced by D. C. Spencer in 1970 ([8] [25] for the linear framework, [9] [48] for the non-linear framework) with totally different operators. For this, if *E* is a vector bundle over the base *X*, we shall introduce the *q*-jet bundle $J_q(E)$ over *X* with (local) sections $\xi_q : (x) \to \left(\xi^k(x), \xi_i^k(x), \xi_{ij}^k(x), \cdots\right)$ transforming like the (local) sections $j_q(\xi) : (x) \to \left(\xi^k(x), \partial_i \xi^k(x), \partial_{ij} \xi^k(x), \cdots\right)$. When $T = T(X)$ is the tangent bundle of *X*, the *Spencer operator* $D : J_{q+1}(E) \to T^* \otimes J_q(T)$ and its extension $D : \wedge^r T^* \otimes J_{q+1}(E) \to \wedge^{r+1} T^* \otimes J_q(E)$ defined by $D\left(\alpha \otimes \xi_{q+1}\right) = d\alpha \otimes \xi_q + (-1)^r \alpha \wedge D\xi_{q+1}$ with $D\xi_{q+1} = j_1\left(\xi_q\right) - \xi_{q+1}$ as we saw for the inverse system, allow comparing these sections by considering the differences $\left(\partial_i \xi^k(x) - \xi_i^k(x), \partial_i \xi_j^k(x) - \xi_{ij}^k(x), \cdots\right)$ and so on. When ω is a nondegenerate metric with Christoffel symbols γ and Levi-Civita isomorphism $j_1(\omega) \simeq (\omega, \gamma)$, we consider the second order involutive system $R_2 \subset J_2(T)$ defined by considering the first order Killing system $\mathcal{L}(\xi)\omega = 0$, adding its first prolongation $\mathcal{L}(\xi)\gamma = 0$ and using ξ_2 instead of $j_2(\xi)$. Looking for the first order generating *compatibility conditions* (CC) $\mathcal{D}_1$ of the corresponding second order operator $\mathcal{D}$ just described, we may then look for the generating CC $\mathcal{D}_2$ of $\mathcal{D}_1$ and so on, exactly like in the differential sequence made successively by the *Killing*, *Riemann*, *Bianchi*, ... operators. We may proceed similarly for the injective operator $T \xrightarrow{j_2} C_0(T) = J_2(T)$, finding successively $C_0(T) \xrightarrow{D_1} C_1(T)$ and $C_1(T) \xrightarrow{D_2} C_2(T)$ induced by *D*. When $n = 2$ and ω is the *Euclide* metric, we have a Lie group of isometries with the 3 infinitesimal generators $\left\{\partial_1, \partial_2, x^1\partial_2 - x^2\partial_1\right\}$. If we now consider the Weyl group defined by $\mathcal{L}(\xi)\omega = A\omega$ with $A = cst$ and $\mathcal{L}(\xi)\gamma = 0$, we have to add the only dilatation $x^1\partial_1 + x^2\partial_2$. Collecting the results and exhibiting the induced kernel upper differential sequence, we get the following commutative *fundamental diagram I* where the upper down arrows are monomorphisms while the lower down arrows are epimorphisms Φ_0, Φ_1, Φ_2 :

$$\begin{array}{ccccccccccccccc}
 & & & & & \overset{j_2}{\rightarrow} & & \overset{D_1}{\rightarrow} & & \overset{D_2}{\rightarrow} & & & & & \\
0 & \rightarrow & \tilde{\Theta} & & \rightarrow & 4 & \rightarrow & 8 & \rightarrow & 4 & \rightarrow 0 & & & & \\
0 & \rightarrow & \Theta & & \overset{j_2}{\rightarrow} & 3 & \overset{D_1}{\rightarrow} & 6 & \overset{D_2}{\rightarrow} & 3 & \rightarrow 0 & & & & \textit{Spencer} \\
 & & & & & \downarrow & & \downarrow & & \downarrow & & & & & \\
0 & \rightarrow & 2 & & \overset{j_2}{\rightarrow} & 12 & \overset{D_1}{\rightarrow} & 16 & \overset{D_2}{\rightarrow} & 6 & \rightarrow 0 & & & & \\
 & & \| & & & \downarrow \Phi_0 & & \downarrow \Phi_1 & & \downarrow \Phi_2 & & & & & \\
0 \rightarrow & \Theta & \rightarrow & 2 & \overset{\mathcal{D}}{\rightarrow} & 9 & \overset{\mathcal{D}_1}{\rightarrow} & 10 & \overset{\mathcal{D}_2}{\rightarrow} & 3 & \rightarrow 0 & & & & \textit{Janet} \\
0 \rightarrow & \tilde{\Theta} & \rightarrow & 2 & \overset{\mathcal{D}}{\rightarrow} & 8 & \overset{\mathcal{D}_1}{\rightarrow} & 8 & \overset{\mathcal{D}_2}{\rightarrow} & 2 & \rightarrow 0 & & & &
\end{array}$$

It follows that "*Spencer and Janet play at see-saw*", the dimension of each *Janet bundle* being decreased by the same amount as the dimension of the corresponding *Spencer bundle* is increased, this number being the number of additional parameters multiplied by $dim\left(\wedge^r T^*\right)$ because:

The linear Spencer Sequence is locally isomorphic to the linear gauge sequence for Lie groups, with the main difference that the group is now acting on the manifold, contrary to the previous situation.

More generally, whenever $R_q \subseteq J_q(E)$ is an involutive system of order q on E, we may define the *Janet bundles* F_r for $r = 0,1,\cdots,n$ by the short exact sequences ([8]):

$$0 \rightarrow \wedge^r T^* \otimes R_q + \delta\left(\wedge^{r-1} T^* \otimes S_{q+1} T^* \otimes E\right) \rightarrow \wedge^r T^* \otimes J_q(E) \rightarrow F_r \rightarrow 0$$

We may pick up a section of F_r, lift it up to a section of $\wedge^r T^* \otimes J_q(E)$ that we may lift up to a section of $\wedge^r T^* \otimes J_{q+1}(E)$ and apply D in order to get a section of $\wedge^{r+1} T^* \otimes J_q(E)$ that we may project onto a section of F_{r+1} in order to construct an operator $\mathcal{D}_{r+1} : F_r \rightarrow F_{r+1}$ generating the CC of $\mathcal{D}_r$ in the canonical *linear Janet sequence*:

$$0 \rightarrow \Theta \rightarrow E \xrightarrow{\mathcal{D}} F_0 \xrightarrow{\mathcal{D}_1} F_1 \xrightarrow{\mathcal{D}_2} \cdots \xrightarrow{\mathcal{D}_n} F_n \rightarrow 0$$

If we have two involutive systems $R_q \subset \hat{R}_q \subset J_q(E)$, *the Janet sequence for* R_q *projects onto the Janet sequence for* $\hat{R}_q$ and we may define inductively *canonical epimorphisms* $F_r \rightarrow \hat{F}_r \rightarrow 0$ for $r = 0,1,\cdots,n$ by comparing the previous sequences for R_q and $\hat{R}_q$, as we already saw.

We can also define the Spencer bundles C_r for $r = 0,1,\cdots,n$ by the short exact sequences ([8]):

$$0 \rightarrow \delta\left(\wedge^{r-1} T^* \otimes g_{q+1}\right) \rightarrow \wedge^r T^* \otimes R_q \rightarrow C_r \rightarrow 0$$

We may pick up a section of C_r, lift it to a section of $\wedge^r T^* \otimes R_q$, lift it up to a section of $\wedge^r T^* \otimes R_{q+1}$ and apply D in order to construct a section of $\wedge^{r+1} \otimes R_q$ that we may project to C_{r+1} in order to construct an operator $D_{r+1} : C_r \rightarrow C_{r+1}$ generating the CC of D_r in the canonical *linear Spencer sequence* which is *another completely different resolution* of the set Θ of (formal) solutions of R_q:

$$0 \rightarrow \Theta \xrightarrow{j_q} C_0 \xrightarrow{D_1} C_1 \xrightarrow{D_2} C_2 \xrightarrow{D_2} \cdots \xrightarrow{D_n} C_n \rightarrow 0$$

However, if we have two systems as above, *the Spencer sequence for* R_q *is now contained into the Spencer sequence for* $\hat{R}_q$ and we may construct inductively *canonical monomorphisms* $0 \to C_r \to \hat{C}_r$ for $r = 0,1,\cdots,n$ by comparing the previous sequences for R_q and $\hat{R}_q$.

When dealing with applications, we have set $E = T$ and considered systems of finite type Lie equations determined by Lie groups of transformations. *In this specific case*, it can be proved that the Janet and Spencer sequences are formally exact, both with their respective adjoint sequences ([11] [28] [32] [36] [38]), namely $ad(\mathcal{D}_r)$ generates the CC of $ad(\mathcal{D}_{r+1})$ while $ad(D_r)$ generates the CC of $ad(D_{r+1})$. We have obtained in particular $C_r = \wedge^r T^* \otimes R_q \subset \wedge^r T^* \otimes \hat{R}_q = \hat{C}_r$ when comparing the classical and conformal Killing systems, but *these bundles have never been used in physics*. Therefore, instead of the classical Killing system $R_2 \subset J_2(T)$ defined by $\Omega \equiv \mathcal{L}(\xi)\omega = 0$ *and* $\Gamma \equiv \mathcal{L}(\xi)\gamma = 0$ or the conformal Killing system $\hat{R}_2 \subset J_2(T)$ defined by $\Omega \equiv \mathcal{L}(\xi)\omega = A(x)\omega$ and $\Gamma \equiv \mathcal{L}(\xi)\gamma = \left(\delta_i^k A_j(x) + \delta_j^k A_i(x) - \omega_{ij}\omega^{ks} A_s(x)\right) \in S_2T^* \otimes T$, we may introduce the *intermediate differential system* $\tilde{R}_2 \subset J_2(T)$ defined by $\mathcal{L}(\xi)\omega = A\omega$ with $A = cst$ and $\Gamma \equiv \mathcal{L}(\xi)\gamma = 0$, for the *Weyl group* obtained by adding the only dilatation with infinitesimal generator $x^i\partial_i$ to the Poincaré group, exactly like we already did when $n = 2$. We have $R_1 \subset \tilde{R}_1 = \hat{R}_1$ but the strict inclusions $R_2 \subset \tilde{R}_2 \subset \hat{R}_2$ and we discover *exactly* the group scheme used through this paper, both with the need to *shift by one step to the left* the physical interpretation of the various differential sequences used. Indeed, as $\hat{g}_2 \simeq T^*$ because $\xi_{ri}^r(x) = nA_i(x)$, the first Spencer operator $\hat{R}_2 \xrightarrow{D_1} T^* \otimes \hat{R}_2$ is induced by the usual Spencer operator $\hat{R}_3 \xrightarrow{D} T^* \otimes \hat{R}_2 : \left(0,0,\xi_{rj}^r, \xi_{rij}^r = 0\right) \to \left(0, \partial_i 0 - \xi_{ri}^r, \partial_i \xi_{rj}^r - 0\right)$ and thus projects by cokernel onto the induced operator $T^* \to T^* \otimes T^*$. Composing with δ, it projects therefore onto $T^* \xrightarrow{d} \wedge^2 T^* : A \to dA = F$ as in EM and so on by using the fact that D_1 *and d are both involutive*, or the composition of epimorphisms:

$$\hat{C}_r \to \hat{C}_r / \tilde{C}_r \simeq \wedge^r T^* \otimes \left(\hat{R}_2 / \tilde{R}_2\right) \simeq \wedge^r T^* \otimes \hat{g}_2 \simeq \wedge^r T^* \otimes T^* \xrightarrow{\delta} \wedge^{r+1} T^*$$

The main result we have obtained is thus to be able to increase the order and dimension of the underlying jet bundles and groups, proving therefore that any 1-form with value in the second order jets $\hat{g}_2$ (*elations*) of the conformal Killing system (conformal group) can be decomposed uniquely into the direct sum (R, F) where R is a section of the *Ricci bundle* S_2T^* and the EM field F is a section of $\wedge^2 T^*$ (Compare to [49]).

Lippmann got the Nobel prize in 1908 for the discovery of color photography. Only one year later, in 1909, the brothers E. and F. Cosserat wrote their "*Théorie des corps déformables*" ([12]) and it is in this book that the previous analogies are quoted for the first time. Between 1895 and 1910, the two brothers published together a series of Notes in the "Comptes Rendus de l'Académie des Sciences de

Paris" and long Notes in famous textbooks or treatises on the mathematical foundations of elasticity theory (Compare to [50]). In particular, they proved that one can exhibit all the concepts and formulas to be found in *elasticity theory* (deformation/strain, compatibility conditions, stress, stress equations, constitutive relations, ...) just by knowing the group of rigid motions of ordinary 3-dimensional space with 3 translations and 3 rotations [10] [51].

It is rather astonishing that *all the formulas* that can be found in the book written by E. and F. Cosserat in 1909 are nothing else but the formal adjoint of the Spencer operator for the Killing equations. More precisely, a section ξ_2 of the first prolongation $R_2 \subset J_2(T)$ of the system $R_1 \subset J_1(T)$ of *Killing equations* is a section of the 2-jet bundle $J_2(T)$ of the tangent bundle $T = T(X)$, namely a set of functions $\xi^k(x), \xi^k_i(x), \xi^k_{ij}(x)$, transforming like the derivatives $\xi^k(x), \partial_i \xi^k(x), \partial_{ij}\xi^k(x)$ of a vector field ξ but also satisfying the linear equations:

$$\omega_{rj}\xi^r_i(x) + \omega_{ir}\xi^r_j(x) + \xi^r(x)\partial_r\omega_{ij} = 0, \quad \xi^k_{ij}(x) = 0$$

where ω is the euclidean metric. Multiplying by test fuctions σ and μ respectively the zero and first order components of the image $D\xi_2$ of the corresponding Spencer operator D, then integrating by part while moving up and down the dumb indices by means of the metric, we successively obtain:

$$\begin{aligned}
&\sigma^i_k\left(\partial_i\xi^k - \xi^k_i\right) + \mu^{j,i}_k\left(\partial_i\xi^k_j - \xi^k_{ij}\right)\\
&= \sigma^{ir}\partial_r\xi_i - \sigma^{ij}\xi_{i,j} + \mu^{ij,r}\partial_r\xi_{i<j}\\
&= -\left[\left(\partial_r\sigma^{ir}\right)\xi_i + \left(\partial_r\mu^{ij,r} + \sigma^{ij} - \sigma^{ji}\right)\xi_{i<j}\right] + divergence\\
&\Rightarrow \partial_r\sigma^{ir} = f^i, \quad \partial_r\mu^{ij,r} + \sigma^{ij} - \sigma^{ji} = m^{ij}
\end{aligned}$$

with evident notations for the Einstein summations involved (Compare to [12] p 137 and 167).

Keeping in mind that, in space-time, there are 4 translations $\left(\xi^k\right)$ and 6 rotations $\left(\xi^k_i\right)$ (3 space rotations + 3 Lorentz transformations), we recover all the $4+6+1+4=15$ variations that can be found in the engineering calculus leading to finite element computations (MODULEF library for example). In addition, we have proved in many books ([10] [11] [18] [28]) and papers ([39] [41] [42]) that the conformal group of space-time is the biggest group of invariance of the Minkowski constitutive laws of EM in vacuum while both sets of Maxwell equations are invariant by any diffeomorphism (*care*!). In particular, considering the space-time dilatation $x^i \to ax^i$ for $i = 1,2,3,4$ with infinitesimal generator $x^i\partial_i$, a transformation which has no intuitive meaning, and gauging the connected compoment $[0,+\infty[$ of the identity with the distinguished identity 1, that is to say transforming the group parameters into functions, just explains why *there must be a zero lower bound in the measure of absolute temperature, both with a distinguished value and invariance under* $T \to 1/T$.

This result clarifies the Helmholtz analogy within jet theory. Indeed, if T is identified with the inverse of a first jet of dilatation, then *T behaves like the derivative of a function without being such a proper derivative*, and we find again exactly the definition of a jet coordinate. *Such a result should lead in the future to revisit the foundations of thermostatics and thermodynamics* ([42]).

The additional 4 transformations, called *elations*, are highly nonlinear and we understand that, contrary to E. and F. Cosserat who succeeded in dealing with the linear transformations, H. Weyl did not succeed in relating electromagnetism with the second order jets of the conformal group in ([13]), *though the idea was a genious one*, simply because he could not use in 1920 a mathematical tool created in 1970 ([25] [47]) but only effective in 1983 ([9] [10] [11] [41] [42]).

The reader may now understand that such a geometric unification was indeed the dream of the brothers E. and F. Cosserat who refer many times explicitly to the work of Mach and Lippmann ([12], p 147, 211). More precisely, using now the *conformal Killing equations*, we have:

$$\omega_{rj}\xi_i^r(x)+\omega_{ir}\xi_j^r(x)+\xi^r(x)\partial_r\omega_{ij}=A(x)\omega_{ij},$$
$$\xi_{ij}^k(x)=\delta_i^kA_j(x)+\delta_j^kA_i(x)-\omega_{ij}\omega^{kr}A_r(x)$$

where $A(x)$ is an arbitrary function and $A_i(x)dx^i$ is an arbitrary 1-form, we get $\xi_{ri}^r(x)=nA_i(x)$ and $\xi_{ijr}^k(x)=0$ for $n\geq 3$ [8] [9] [10] [11]. Accordingly, the zero, first and second order components (*field*) of the image $D\xi_3$ of the Spencer operator D are:

$$\partial_i\xi^k-\xi_i^k,\quad \partial_i\xi_j^k-\xi_{ij}^k,\quad \partial_r\xi_{ij}^k-\xi_{ijr}^k=\partial_r\xi_{ij}^k\Rightarrow\partial_i\xi_{rj}^r-\xi_{rij}^r=\partial_i\xi_{rj}^r$$

and we can recover $\epsilon_{ij}=\frac{1}{2}\left[\left(\partial_i\xi_j-\xi_{i,j}\right)+\left(\partial_j\xi_i-\xi_{j,i}\right)\right]$. Identifying the speed with a (gauged) Lorentz rotation, that is to say setting $\partial_4\xi^k-\xi_4^k=0$ as a constraint ([12]), we can therefore measure both $\partial_4\xi_4^k-\xi_{44}^k=\left(1/c^2\right)\gamma^k-\omega^{kr}A_r$ for $k=1,2,3$ (care to the sign!) and $\partial_i\xi_r^r-\xi_{ri}^r=-n\left(\left(1/T\right)\partial_iT+A_i\right)$, thus $\left(1/T\right)\nabla T+\left(1/c^2\right)\boldsymbol{\gamma}$ by substraction, where $\boldsymbol{\gamma}$ is the acceleration, and thus $(1/T)\nabla T$ in first approximation ([52], p. 922). Also, the formula $\partial_i\xi_{rj}^r-\partial_j\xi_{ri}^r=n\left(\partial_iA_j-\partial_jA_i\right)=nF_{ij}$ exactly describes the results of [13] by means of the Spencer operator and explains why the EM field is on equal footing with deformation and gradient of temperature, contrary to its status in gauge theory.

Roughly speaking, E. and F. Cosserat were only using the zero and first order components of the image of the Spencer operator while H. Weyl was only using the first and second order components (See [38] for more comments and [9] [10] [11] [39] [42] for a nonlinear version).

4. Conclusions

Recapitulating all the results previously obtained, we may finally say (See the end

of [39]):

- "*Beyond the mirror*" of the classical approach to apparently well known and established theories, there is a totally new interpretation of these theories and the corresponding field/matter couplings by means of the Spencer sequence for the conformal Killing operator.
- The purely mathematical results of Section 3 perfectly agree with the origin and existence of elastic and electromagnetic waves but question the origin and existence of gravitational waves because the parametrization of the *Cauchy* operator can be simply done by the adjoint of the *Ricci* operator without any reference to the *Einstein* or even *Bianchi* operators. We believe that such a confusion mainly came from the fact that it had never been noticed that the *Einstein* operator was self-adjoint.
- They prove that the concept of "*field*" in a physical theory must not be related with the concept of "*curvature*" because it is a 1-form with value in a Lie algebroid (first Spencer bundle) and *not* a 2 form with value in a Lie algebra (second Spencer bundle). The "*shift by one step*" in the physical interpretation of a differential sequence is thus the main feature of this new mathematical framework.
- They also prove that gravitation and electromagnetism have a common conformal origin. In particular, electromagnetism has only to do with the conformal group of space-time and not with $U(1)$ as it is still believed today in Gauge Theory.

Conflicts of Interest

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

References

[1] de Broglie, L. (1964) Thermodynamique de la Particule isolée, Gauthiers-Villars, Paris.

[2] Mach, E. (1872) Die Geschichte und die Wurzel des Satzes von der Erhaltung der Arbeit. Calve, Prag, 54.

[3] Lippmann, G. (1876) *Comptes rendus de l'Académie des Sciences*, **82**, 1425-1428.

[4] Mach, E. (1900) Prinzipien der Wärmelehre. J. A. Barth, Leipzig, 330.

[5] Lippmann, G. (1907) *Ann. Phys. Chem.*, **23**, 994-996.

[6] Ougarov, V. (1969) Théorie de la Relativité Restreinte. MIR, Moscow (French Translation, 1979).

[7] Vessiot, E. (1903) *Annales Scientifiques de l'École Normale Supérieure*, **20**, 411-451. http://numdam.org
https://doi.org/10.24033/asens.529

[8] Pommaret, J.-F. (1978) Systems of Partial Differential Equations and Lie Pseudogroups. Gordon and Breach, New York; Russian Translation: MIR, Moscow (1983).

[9] Pommaret, J.-F. (1983) Differential Galois Theory. Gordon and Breach, New

York.

[10] Pommaret, J.-F. (1988) Lie Pseudogroups and Mechanics. Gordon and Breach, New York.

[11] Pommaret, J.-F. (1994) Partial Differential Equations and Group Theory. Kluwer, Dordrecht. https://doi.org/10.1007/978-94-017-2539-2

[12] Cosserat, E. and Cosserat, F. (1909) Théorie des Corps Déformables. Hermann, Paris.

[13] Weyl, H. (1918) Space, Time, Matter. Springer, London.

[14] Pommaret, J.-F. (2001) *Acta Mechanica*, **149**, 23-39. https://doi.org/10.1007/BF01261661

[15] Bourbaki, N. (1980) Algèbre, Chapter 10, Algèbre Homologique. Masson, Paris.

[16] Hu, S.-T. (1968) Introduction to Homological Algebra, Holden-Day.

[17] Northcott, D.G. (1966) An Introduction to Homological Algebra. Cambridge University Press, Cambridge.

[18] Pommaret, J.-F. (2001) Partial Differential Control Theory. Kluwer, Dordrecht.

[19] Pommaret, J.-F. (2005) Algebraic Analysis of Control Systems Defined by Partial Differential Equations. In: Lamnabhi-Lagarrigue, F., Loría, A. and Panteley, E., Eds., *Advanced Topics in Control Systems Theory. Lecture Notes in Control and Information Science*, Volume 311, Chapter 5, Springer, London, 155-223. https://doi.org/10.1007/11334774_5

[20] Rotman, J.J. (1979) An Introduction to Homological Algebra, Pure and Applied Mathematics. Academic Press, London.

[21] Bjork, J.E. (1993) Analytic *D*-Modules and Applications. Kluwer, Dordrecht. https://doi.org/10.1007/978-94-017-0717-6

[22] Kashiwara, M. (1995) Algebraic Study of Systems of Partial Differential Equations. Mémoires de la Société Mathématique de France, 63 (Transl. from Japanese of His 1970 Master's Thesis).

[23] Quadrat, A. and Robertz, R. (2014) *Acta Applicandae Mathematicae*, **133**, 187-234. http://hal-supelec.archives-ouvertes.fr/hal-00925533 https://doi.org/10.1007/s10440-013-9864-x

[24] Schneiders, J.-P. (1994) *Bulletin de la Societe Royale des Sciences de Liege*, **63**, 223-295.

[25] Spencer, D.C. (1965) *Bulletin of the American Mathematical Society*, **75**, 1-114.

[26] Chyzak, F., Quadrat, A. and Robertz, D. (2007) OreModules: A Symbolic Package for the Study of Multidimensional Linear Systems. In: Chiasson, J. and Loiseau, J.J., Eds., *Applications of Time Delay Systems. Lecture Notes in Control and Information Sciences*, Volume 352, Springer, Berlin, Heidelberg, 233-264. http://wwwb.math.rwth-aachen.de/OreModules https://doi.org/10.1007/978-3-540-49556-7_15

[27] Pommaret, J.-F. (2015) *Multidimensional Systems and Signal Processing*, **26**, 405-437. https://doi.org/10.1007/s11045-013-0265-0

[28] Pommaret, J.-F. (2016) Deformation Theory of Algebraic and Geometric Structures. Lambert Academic Publisher, Saarbrücken, Germany. In: *Topics in Invariant Theory. Lecture Notes in Mathematics*, Volume 1478, Springer, Berlin, Heidelberg, 244-254. http://arxiv.org/abs/1207.1964 https://doi.org/10.1007/BFb0083506

[29] Pommaret, J.-F. (2018) New Mathematical Methods for Physics. In: *Mathematical Physics Books*, Nova Science Publishers, New York.

[30] Oberst, U. (1990) *Acta Applicandae Mathematica*, **20**, 1-175. https://doi.org/10.1007/BF00046908

[31] Zerz, E. (2000) Topics in Multidimensional Linear Systems Theory. In: *Lecture Notes in Control and Information Sciences*, *LNCIS* 256, Springer, Heidelberg.

[32] Pommaret, J.-F. (2017) *Journal of Modern Physics*, **8**, 2122-2158. https://doi.org/10.4236/jmp.2017.813130

[33] Airy, G.B. (1863) *Philosophical Transactions of the Royal Society of London*, **153**, 49-80. https://doi.org/10.1098/rstl.1863.0004

[34] Beltrami, E. (1892) Osservazioni sulla Nota Precedente. *Atti della Accademia Nazionale dei Lincei*, **5**, 141-142.

[35] Maxwell, J.C. (1870) *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, **26**, 1-40. https://doi.org/10.1017/S0080456800026351

[36] Pommaret, J.-F. (2016) *Journal of Modern Physics*, **7**, 699-728. https://doi.org/10.4236/jmp.2016.77068

[37] Pommaret, J.-F. (2019) *Journal of modern Physics*, **10**, 1454-1486 https://doi.org/10.4236/jmp.2019.1012097

[38] Pommaret, J.-F. (2010) *Acta Mechanica*, **215**, 43-55. https://doi.org/10.1007/s00707-010-0292-y

[39] Pommaret, J.-F. (2012) Spencer Operator and Applications: From Continuum Mechanics to Mathematical Physics. In: Gan, Y., Ed., *Continuum Mechanics-Progress in Fundamentals and Engineering Applications*, InTech, London. https://doi.org/10.5772/35607

[40] Pommaret, J.-F. (2013) *Journal of Modern Physics*, **4**, 223-239. https://doi.org/10.4236/jmp.2013.48A022

[41] Pommaret, J.-F. (2014) *Journal of Modern Physics*, **5**, 157-170. https://doi.org/10.4236/jmp.2014.55026

[42] Pommaret, J.-F. (2015) From Thermodynamics to Gauge Theory: The Virial Theorem Revisited. In: *Gauge Theories and Differential Geometry*, NOVA Science Publisher, New York, 1-46.

[43] Foster, J. and Nightingale, J.D. (1979) A Short Course in General Relativity. Longman, London.

[44] Choquet-Bruhat, Y. (2015) Introduction to General Relativity, Black Holes and Cosmology. Oxford University Press, Oxford.

[45] Poincare, H. (1901) *Comptes rendus de l'Académie des Sciences*, **132**, 369-371.

[46] Arnold, V. (1974) Méthodes Mathématiques de la Mécanique Classique, Apendice 2.

[47] Janet, M. (1920) *Journal de Mathématiques*, **8**, 65-151.

[48] Kumpera, A. and Spencer, D.C. (1972) Lie Equations. In: *Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ. https://doi.org/10.1515/9781400881734

[49] Zou, Z., Huang, P., Zang, Y. and Li, G. (1979) *Scientia Sinica*, **22**, 628-636.

[50] Teodorescu, P.P. (1975) Dynamics of Linear Elastic Bodies. Abacus Press, Tunbridge, Wells (Editura Academiei, Bucuresti, Romania).

[51] Pommaret, J.-F. (1997) *Annales des Ponts et Chaussées*, **82**, 59-66. (Translation by D.H. Delphenich)

[52] Eckart, C. (1940) *Physical Review Journals Archive*, **58**, 919-924. https://doi.org/10.1103/PhysRev.58.919

Main Mathematical Notations

X manifold with tangent, cotangent, symmetric, skewsymmetric bundles $T, T^*, S_qT^*, \wedge^r T^*$

$J_q(E)$ q-jet bundle of the vector bundle E over X with $dim(X)=n$

$R_q \subset J_q(E)$ involutive system of order q on E with symbol $g_q = R_q \cap S_qT^* \otimes E$

$F_r = \wedge^r T^* \otimes J_q(E) \Big/ \left(\wedge^r T^* \otimes R_q + \delta\left(\wedge^{r-1} T^* \otimes S_{q+1} T^* \otimes E \right) \right)$ Janet bundles

$C_r(E) = \wedge^r T^* \otimes J_q(E) \Big/ \delta\left(\wedge^{r-1} T^* \otimes S_{q+1} T^* \otimes E \right)$ generic Spencer bundles

$C_r = \wedge^r T^* \otimes R_q \Big/ \delta\left(\wedge^{r-1} T^* \otimes g_{q+1} \right) \subset C_r(E)$ Spencer bundles

$0 \to C_r \to C_r(E) \xrightarrow{\Phi_r} F_r \to 0$ exact sequences with Φ_r induced by $\Phi = \Phi_0$ and $R_q = ker(\Phi)$

$C_0 \xrightarrow{D_1} C_1 \xrightarrow{D_2} \cdots \xrightarrow{D_n} C_n \to 0$ Spencer sequence with D_r induced by $D: R_{q+1} \to T^* \otimes R_q$

$E \xrightarrow{\mathcal{D}} F_0 \xrightarrow{\mathcal{D}_1} F_1 \xrightarrow{\mathcal{D}_2} \cdots \xrightarrow{\mathcal{D}_n} F_n \to 0$ Janet sequence of successive compatibility conditions (CC)

G Lie group with identity e and Lie algebra $\mathcal{G} = T_e(G)$ with bracket []

$X \times G \to T^* \otimes \mathcal{G} \xrightarrow{MC} \wedge^2 T^* \otimes \mathcal{G}$ Maurer-Cartan nonlinear gauge sequence

$a \to a^{-1}da = A \to dA - [A, A] = F$ pull back of $T(a): T(X) \to T(G)$ and Cartan curvature.

Journal of Modern Physics, 2019, 10, 1596-1605
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Influence of Both Magnetic Field and Temperature on Silicon Solar Cell Base Optimum Thickness Determination

Nouh Mohamed Moctar Ould Mohamed[1], Ousmane Sow[2], Sega Gueye[1], Youssou Traore[1], Ibrahima Diatta[1], Amary Thiam[1], Mamour Amadou Ba[1], Richard Mane[1], Ibrahima Ly[3], Gregoire Sissoko[1]

[1]Laboratory of Semiconductors and Solar Energy, Physics Department, Faculty of Science and Technology, University Cheikh Anta Diop, Dakar, Senegal
[2]University Institute of Technology, University of Thiès, Thiès, Sénégal
[3]Ecole Polytechnique de Thiès, Thiès, Sénégal
Email: gsissoko@yahoo.com

How to cite this paper: Mohamed, N.M.M.O., Sow, O., Gueye, S., Traore, Y., Diatta, I., Thiam, A., Ba, M.A., Mane, R., Ly, I. and Sissoko, G. (2019) Influence of Both Magnetic Field and Temperature on Silicon Solar Cell Base Optimum Thickness Determination. *Journal of Modern Physics*, **10**, 1596-1605.
https://doi.org/10.4236/jmp.2019.1013105

Received: October 29, 2019
Accepted: November 22, 2019
Published: November 25, 2019

Abstract

The minority carrier's recombination velocity at the junction and at the back surface is used for the modeling and determination of the optimum thickness of the base of a silicon solar cell in the static regime, under magnetic field and temperature influence. This study takes into account the Umklapp process and the Lorentz effect on the minority carriers photogenerated in the base.

Keywords

Silicon Solar Cell, Diffusion Coefficient, Surface Recombination Velocity, Optimum Base Thickness, Lorentz and Umklapp Processes

1. Introduction

The parameters of the electric equivalent model of the solar cell, under darkness or illumination lead to its characterization, through the measurement of electrical current and voltage (I-V) [1].

These measurements are made by maintaining the solar cell under static [2] [3] or dynamic (transient [4] [5] [6] or frequency [7] [8] [9]) regimes and thus defining the different characterization technics.

From these technics, the series (Rs) [10] [11] [12] and shunt (Rsh) [13] [14] [15] resistances are deduced, which allow appreciating the quality of the solar cell, as well as the capacitance of transition.

However, these parameters are not expressed as function of solar cell geometrical parameters, especially the thicknesses of emitter, space charge region, base, and grain size [16].

In this work, the phenomenological parameters, such as minority carrier recombination velocity [17] [18] in the bulk (τ) [19], at the emitter-base junction (Sf) and at the back surface (Sb) [20] [21] of the base with thickness (H), are studied to extract the optimum thickness H, of the base of the silicon solar cell, placed under magnetic field B and with variation of the temperature T. This optimum thickness is obtained from the curves of the back surface recombination velocity, expressed as dependent of both the diffusion coefficient $D(T, B)$ [22] [23] revealing the Umklapp process [24] and the base thickness.

2. Theory

Figure 1 represents a n^+-p-p^+ silicon solar cell under polychromatic illumination, by the emitter (n^+), through the collected grids. The space charge region (SCR), in $x = 0$, constitutes the junction (n^+-p), allowing the separation of photogenerated electron-hole pairs. The rear face (p^+), in $x = H$, is a zone where an electric field exists (Back surface Field), which allows the return of the minority charge carriers towards the junction [25] [26].

When the solar cell is under illumination, the density $\delta(x, T, B)$ of photogenerated carriers in the base under magnetic field B at temperature T, is governed by the following magneto transport equation.

$$D(T,B)\frac{\partial^2\delta(x,T,B)}{\partial x^2}-\frac{\delta(x,T,B)}{\tau}=-G(x) \tag{1}$$

τ et $D(T, B)$ are, respectively, the lifetime and the diffusion coefficient of the excess minority carriers in the base, under magnetic field and under temperature.

Under magnetic field, the diffusion coefficient is given by the following relation [27]:

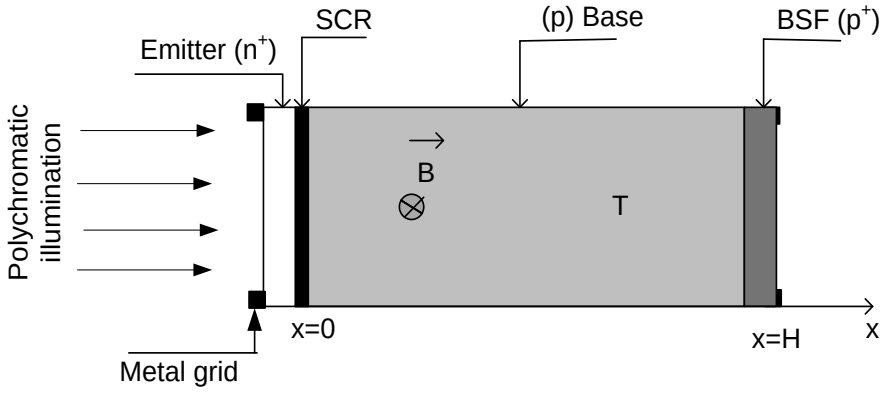


Figure 1. The solar cell structure to the n^+-p-p^+ type, under both magnetic field and temperature.

$$D(T,B)=\frac{D(T)}{1+(\mu B)^2} \tag{2}$$

With

$$D(T)=\mu(T)\cdot\frac{K_b\cdot T}{q} \tag{3}$$

And the mobility coefficient is given as [28]:

$$\mu(T)=1.43*10^9\cdot T^{-2.42} \tag{4}$$

$L^2(T,B)=D(T,B)$. L represents the diffusion length of minority carriers in excess. Plot of expression (2) allows to extract maximum diffusion coefficient which is related to optimum temperature by following relation [23]:

$$D_{\max}(B,T)=\alpha'\cdot\left[T_{op}(B)\right]^{\beta'} \tag{5}$$

With

$$\alpha'=-1.51\ \mathrm{cm^2/s\cdot K} \tag{6}$$

and

$$\beta'=11.87 \tag{7}$$

$\delta(x,T,B)$ the density of photogenerated carriers in the base, is produced by the generation rate, expressed by [29]:

$$G(x)=n\sum_{i=1}^{3}a_i\mathrm{e}^{-b_i x} \tag{8}$$

where n is the number of sun or level of illumination.

The parameters a_i et b_i stem from the modeling of the incident illumination as defined under A.M1.5.

The expression of the excess minority carrier density in the base is given by the resolution of the continuity equation and is written as:

$$\begin{aligned}\delta(x,Sf,n,H,T,B)=&A(Sf,n,H,T,B)\cdot\cosh\left(\frac{x}{L(T,B)}\right)\\&+E(Sf,n,H,T,B)\cdot\sinh\left(\frac{x}{L(T,B)}\right)\\&+\sum_{i=1}^{3}K(n,T,B)\cdot\mathrm{e}^{-b_i x}\end{aligned} \tag{9}$$

where

$$K(n,T,B)=\frac{na_i\cdot L^2(T,B)}{D(T,B)\left[1-\left(L(T,B)\cdot b_i\right)^2\right]} \tag{10}$$

$A(Sf,n,H,T,B)$ and $E(Sf,n,H,T,B)$ are coefficients determined from the boundary conditions which respectively introduce the surface recombination velocity at the junction (*Sf*) and at the rear face (*Sb*).

1) At the junction $x=0$ (SCR)

$$D(T,B)\frac{\partial\delta(x,Sf,n,H,T,B)}{\partial x}\bigg|_{x=0} = Sf*\delta(0,Sf,n,H,T,B) \quad (11)$$

Sf indicates the velocity of passage of the charge carriers across the junction, to the emitter. This velocity of passage of the minority carriers is governed by the external charge resistance conncted to the solar cell which imposes the operating point [19] [20] [21]. Thus, the carriers that have crossed through the junction and are not collected, constitute the losses through the shunt resistance, associated with the intrinsic minority carriers velocity *S*0, at the junction [4] [9] [20].

2) At the back surface $x = H(BSF)$

$$D(T,B)\cdot\frac{\partial\delta(x,Sf,n,H,T,B)}{\partial x}\bigg|_{x=H} = -Sb*\delta(Sf,n,H,T,B) \quad (12)$$

H is the thickness of the base of the solar cell.

Sb is the back surface recombination velocity of the excess minority carrier [20] [21] [25] [30], at coordinate $x = H$, where exists a rear electric field (p/p^+, low-high junction), which returns the carriers toward the junction (SCR), to be collected.

3. Results and Discussions

3.1. Photocurrent Density

Fick's law allows us to obtain the expression of the photocurrent density. This expression is given by the following equation:

$$\begin{aligned} Jph(x,Sf,n,H,T,B) &= qD(T,B)\cdot\frac{\partial\delta(x,Sf,n,H,T,B)}{\partial x}\bigg|_{x=0} \\ &= qD(T,B)\cdot Sf\cdot\delta(0,Sf,n,H,T,B) \end{aligned} \quad (13)$$

Figure 2 gives the profile of the photocurrent density versus minority carrier recombination velocity at the junction for different values of both, temperature and magnetic field according to Equation (5).

For values of $Sf < 10^2$ cm/s, the photocurrent is practically nil (whatever the temperature and magnetic field), which corresponds to an open circuit operating point of the solar cell. Between 10^4 cm/s and 10^5 cm/s, the photocurrent is increasing.

For *Sf* beyond 10^5 cm/s, the photocurrent is constant with *Sf* and corresponds to the short-circuit current density Jphsc, which is a plateau that increases with the temperature and decreases with magnetic field (Lorentz's law).

3.2. Back Surface Recombination Velocity *Sb* (*T*, *B*)

Figure 2 indicates a plateau regardless of *Sf*, so the derivative of the expression of the photocurrent density with respect to the recombination velocity, is zero. [17] [20] [21] and is written

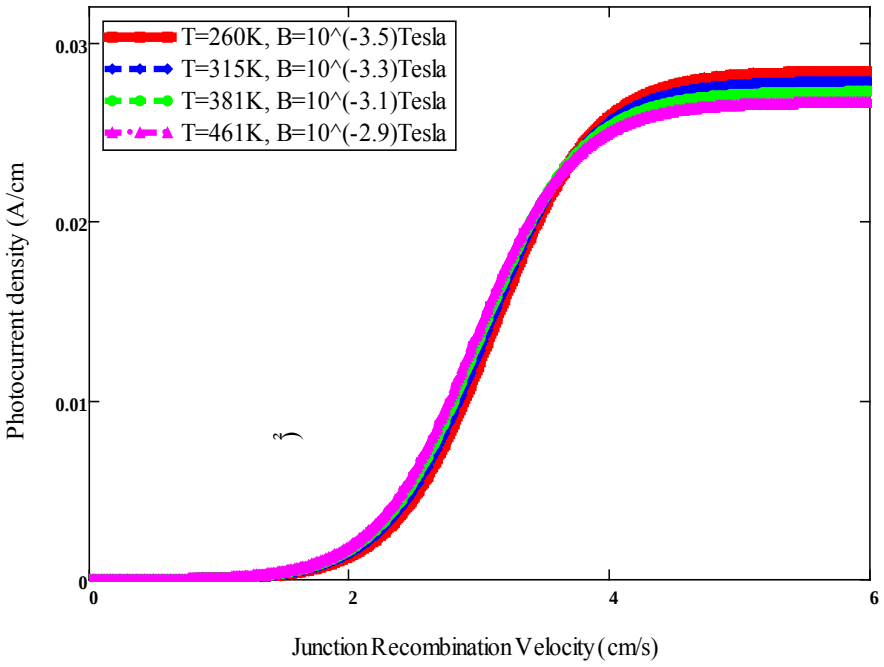


Figure 2. Photocurrent density calibration curves versus minority carrier recombination velocity at the junction under temperature and magnetic field.

$$\frac{\partial Jph(Sf,n,H,T,B)}{\partial Sf}=0 \tag{16}$$

The resolution of this equation leads to following expressions of back surface recombination velocity $Sb1$ (H, D) and $Sb2$ (H, D):

$$Sb1(H,T,B)=\sum_{i=1}^{3}\frac{D(T,B)*\left[b_i\left(\cosh\left(\frac{H}{L(T,B)}\right)-e^{-b_iH}\right)-\frac{1}{L(T,B)}\sinh\left(\frac{H}{L(T,B)}\right)\right]}{\cosh\left(\frac{H}{L(T,B)}\right)-e^{-b_iH}-L(T,B)b_i\sinh\left(\frac{H}{L(T,B)}\right)} \tag{17}$$

where in $Sb1$ expression, appears the effect of light absorption (b_i coefficient) in the material and leads to a generation rate ($b_i * D(B,T)$, for $L >> H$) [20] [21], [23]. Given Fick's law on the back surface, the $Sb2$ (<0) expression is the intrinsic minority carriers recombination velocity at the back surface [22].

$$Sb2(H,T,B)=-\frac{D(T,B)}{L(T,B)}*\tanh\left(\frac{H}{L(T,B)}\right) \tag{18}$$

$Sb1$ and $Sb2$ lead asymptotically to the quotient D/L, which is the diffusion velocity (for $L << H$) [20], [21].

Figure 3 gives the profile of the two expressions of minority carrier recombination velocity at the rear face versus the solar cell base thickness, for different diffusion coefficient $D_{max}(B, T)$ values of the minority carriers in the base undergoing the Umklapp (thermal agitation) and Lorent processes(defection) [23] [24].

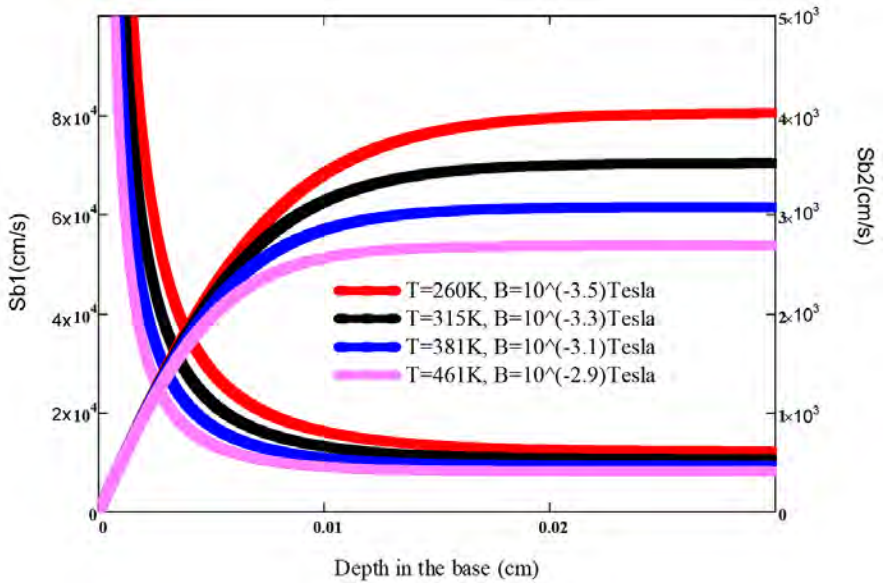


Figure 3. Back surface recombination velocity curves versus solar cell base thickness.

The intersection of the curves *Sb*1 and *Sb*2, has for abscissa, the optimum thickness of the base of the solar cell for each diffusion coefficient $D_{max}(B, Topt)$ [23].

Table 1 summarizes the variation of the solar cell base thickness for each diffusion coefficient ($D_{max}(B, Topt)$) and the respective short-circuit currents Jsc1 and Jsc2 which remain maximum and constant.

Figure 4 and **Figure 5**, give the representation of the thickness of the base of the solar cell necessary for each case of the diffusion coefficient. The correlation between the diffusion coefficient $D_{max}(T, B)$ and the optimum base thickness is established.

The current-voltage characteristics, under constant illumination of the solar cell having different base thicknesses, are simulated and the efficiency is obtained according to the thickness and under the influence of the surface recombination velocity [31]. The influence of thickness is highlighted in dynamic regime [16] [30] [32] through the constant decay time, as well as in studies of the solar cell in 3D model [24] [33] [34] where the electrical (*D*, *Sf*, *Sb*) [35] [36] and geometry (grain size) [37] parameters are involved.

Thus the results we propose in this work, constitute a contribution for the modelling and manufacturing of the solar cell thickness, for optimum efficiency under specific operating conditions [36].

4. Conclusion

The proposed study on determining the optimum thickness of the base of a silicon solar cell under temperature and magnetic field, takes into account the behaviour of minority carriers in physical processes of thermal agitation (Umklapp) and deflection (Lorentz). These physical mechanisms are quantified through the diffusion coefficient and recombination velocity of minority carriers

Table 1. Base optimum thickness Value (H) for different diffusion coefficients $D_{max}(B, Topt)$.

(T(K), B(Tesla))	(261; $10^{-3.5}$)	(315; $10^{-3.3}$)	(381; $10^{-3.1}$)	(461; $10^{-2.9}$)
D_{max} (cm^2/s)	32.348	24.689	18.843	14.381
H (cm)	0.00821	0.00721	0.00633	0.00551
$Jsc1$ (A/cm^2)	0.027	0.026	0.026	0.025
$Jsc2$ (A/cm^2)	0.028	0.027	0.027	0.026
$Sb1$ (cm/s)	18584	16186	14261	12435
$Sb2$ (cm/s)	3097.3	2679.7	2376.9	2072.5

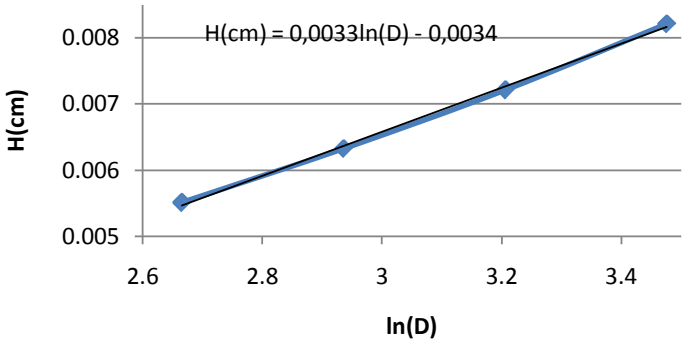


Figure 4. Base depth H as function of logarithm of diffusion coefficient.

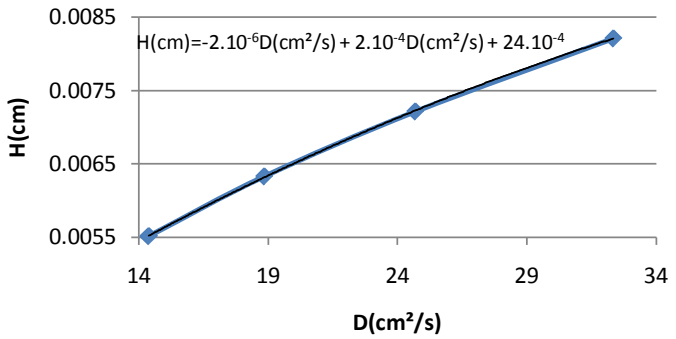


Figure 5. Base depth H as function of diffusion coefficient.

at the back surface. Thus, the study of the photocurrent obtained from the solar cell in short-circuit operation, allows extracting the theoretical expressions of the recombination velocity at the back surface of the base. The optimum thickness is then deduced through the analysis of the profile of these two expressions of minority carrier recombination velocity versus thickness, represented by the intercept points. The optimum thickness required for the manufacturing of the base is then established. Thus, solar cell can operate under (T, B) conditions and yield optimum efficiency.

Conflicts of Interest

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

References

[1] Sow, O., Zerbo, I., Mbodji, S., Ngom, M.I., Diouf, M.S. and Sissoko, G. (2012) *International Journal of Science*, *Environment and Technology*, **1**, 230-246.

[2] Diasse, O., Diao, A., Wade, M., Diouf, M., Diatta, I., Mane, R., Traore, Y. and Sissoko, G. (2018) *Journal of Modern Physics*, **9**, 189-201. https://doi.org/10.4236/jmp.2018.92012

[3] Barro, F.I., Gaye, S., Deme, M., Diallo, H.L., Samb, M.L., Samoura, A.M., Mbodji, S. and Sissoko, G. (2008) Influence of Grain Size and Grain Boundary Recombination Velocity on the Series and Shunt Resistances of a Poly-Crystalline Silicon Solar Cell. *Proceedings of the* 23*rd European Photovoltaic Solar Energy Conference*, Valencia, 1-5 September 2008, 612-615.

[4] Kalidou Mamadou, S.Y., Diene, A., Tamba, S., Diouf, M.S., Diatta, I., Dieye, M., Traore, Y. and Sissoko, G. (2016) *Journal of Scientific and Engineering Research*, **3**, 433-445.

[5] Gaye, I., Corréa, A., Ba, B., Ndiaye, A.L., Nanéma, E., Ba, A.B.B., Adj, M. and Sissoko, G. (1996) *Renewable Energy*, **3**, 1598-1601.

[6] Corréa, A., Gaye, I., Ba, B., Ndiaye, A.L. and Sissoko, G. (1994) *Renewable Energy*, **5**, 166-168. https://doi.org/10.1016/0960-1481(94)90366-2

[7] Traore, Y., Thiam, N., Thiame, M., Thiam, A., Ba, M.L., Diouf, M.S., Diatta, I., Mballo, O., Sow, E.H., Wade, M. and Sissoko, G. (2019) *Journal of Modern Physics*, **10**, 1235-1246. http://www.scirp.org/journal/jmp https://doi.org/10.4236/jmp.2019.1010082

[8] Gueye, M., Diallo, H.L., Moustapha, A.K.M., Traore, Y., Diatta, I. and Sissoko, G. (2018) *World Journal of Condensed Matter Physics*, **8**, 185-196. http://www.scirp.org/journal/wjcmp https://doi.org/10.4236/wjcmp.2018.84013

[9] Thiam, N., Diao, A., Ndiaye, M., Dieng, A., Thiam, A., Sarr, M., Maiga, A.S. and Sissoko, G. (2012) *Research Journal of Applied Sciences*, *Engineering and Technology*, **3**, 602-611.

[10] Bouzidi, K., Chegaar, M. and Bouhemadou, A. (2007) *Solar Energy Materials & Solar Cells*, **91**, 1647-1651. https://doi.org/10.1016/j.solmat.2007.05.019

[11] El-Adawi, M.K. and Al-Nuaim, I.A. (2002) *Vacuum*, **64**, 33-36. https://doi.org/10.1016/S0042-207X(01)00370-0

[12] Bashahu, M. and Habyarimana, A. (1995) *Renewable Energy*, **6**, 127-138. https://doi.org/10.1016/0960-1481(94)E0021-V

[13] Diouf, M.S., Sahin, G., Thiam, A., Faye, K., Ngom, M.I., Gaye, D. and Sissoko, G. (2015) *International Journal of Innovative Science Engineering & Technology*, **2**, 931-938.

[14] Ly, I., Ndiaye, M., Wade, M., Thiam, N., Gueye, S. and Sissoko, G. (2013) *Research Journal of Applied Sciences*, *Engineering and Technology*, **5**, 203-208. https://doi.org/10.19026/rjaset.5.5105

[15] Dione, M.M., Diallo, H.L., Wade, M., Ly, I., Thiame, M., Toure, F., Camara, A.G., Dieme, N., Bako, Z.N., Mbodji, S., Barro, F.I. and Sissoko, G. (2011) Determination of the Shunt and Series Resistances of a Vertical Multijunction Solar Cell under Constant Multispectral Light. *Proceedings of the* 26*th European Photovoltaic Solar Energy Conference*, Hamburg, 4-8 September 2011, 250-254.

[16] Honma, N. and Munakata, C. (1987) *Japanese Journal of Applied Physics*, **23**,

2033-2036. https://doi.org/10.1143/JJAP.26.2033

[17] Ly Diallo, H., Dieng, B., Ly, I., Dione, M.M., Ndiaye, M., Lemrabott, O.H., Bako, Z.N., Wereme, A. and Sissoko, G. (2012) *Research Journal of Applied Sciences, Engineering and Technology*, **4**, 2626-2631. http://www.maxwell.org

[18] Dhariwal, S.R. and Vasu, N.K. (1981) *Solid-State Electronics*, **24**, 915-927. https://doi.org/10.1016/0038-1101(81)90112-X

[19] Sissoko, G., Sivoththanam, S., Rodot, M. and Mialhe, P. (1992) Constant Illumination-Induced Open Circuit Voltage Decay (CIOCVD) Method, as Applied to High Efficiency S_i Solar Cells for Bulk and Back Surface Characterization. 11*th European Photovoltaic Solar Energy Conference and Exhibition*, Montreux, 352-354.

[20] Sissoko, G., Corréa, A., Naména, E., Adj, M. and Ndiaye, A.L. ((1998) Recombination Parameters Measurement in Double Sided Surface Field Solar Cell. *World Renewable Energy Congress*, Florence, 20-25 September 1998, 1856-1859.

[21] Bocande, Y.L.B., Correa, A., Gaye, I., Sow, M.L. and Sissoko, G. (1994) *Renewable Energy*, **5**, 1698-1700.

[22] Sissoko, G., Museruka, C., Corréa, A., Gaye, I. and Ndiaye, A.L. (1996) Light Spectral Effect on Recombination Parameters of Silicon Solar Cell. *World Renewable Energy Congress*, Part III, 1487-1490.

[23] Mane, R., Ly, I., Wade, M., Datta, I., Douf, M.S., Traore, Y., Ndiaye, M., Tamba, S. and Sissoko, G. (2017) *Energy and Power Engineering*, **9**, 1-10. https://doi.org/10.4236/epe.2017.91001

[24] Berman, R. (1951) *Nature*, **168**, 277-280. https://doi.org/10.1038/168277a0

[25] Le Quang, N., Rodot, M., Nijs, J., Ghannam, M. and Coppye, J. (1992) *Journal of Physics*, **3**, 1305-1316.

[26] Fossum, G., Burgess, E.L. and Lindholm, F.A. (1978) *Solid-State Electronics*, **27**, 729-737. https://doi.org/10.1016/0038-1101(78)90005-9

[27] Betser, Y., Ritter, D., Bahir, G., Cohen, S. and Serling, J. (1995) *Applied Physics Letters*, **67**, 1883-1884. https://doi.org/10.1063/1.114364

[28] Kunst, M. and Sanders, A. (1992) *Semiconductor Science and Technology*, **7**, 51-59. https://doi.org/10.1088/0268-1242/7/1/009

[29] Furlan, J. and Amon, S. (1985) *Solid-State Electronics*, **28**, 1241-1243. https://doi.org/10.1016/0038-1101(85)90048-6

[30] Joardar, K., Dondero, R.C. and Schroder, D.K. (1989) *Solid-State Electronics*, **32**, 479-483. https://doi.org/10.1016/0038-1101(89)90030-0

[31] Demesmaeker, E., Symons, J., Nijs, J. and Mertens, R. (1991) The Influence of Surface Recombination on the Limiting Efficiency and Optimum Thickness of Silicon Solar Cells. 10*th European Photovoltaic Solar Energy Conference*, Lisbon, 66-67. https://doi.org/10.1007/978-94-011-3622-8_17

[32] Sproul, A.B. (1994) *Journal of Applied Physics*, **76**, 2851-2854. https://doi.org/10.1063/1.357521

[33] Ba, B., Kane, M., Fickou, A. and Sissoko, G. (1993) *Solar Energy Materials and Solar Cells*, **31**, 33-49. https://doi.org/10.1016/0927-0248(93)90005-N

[34] Diallo, H.L., Maiga, A.S., Wereme, A. and Sissoko, G. (2008) *The European Physical Journal Applied Physics*, **42**, 193-211. https://doi.org/10.1051/epjap:2008085

[35] Diop, M., Ba, H., Thiam, N., Diatta, I., Traore, Y., Ba, M., Sow, E., Mballo, O. and Sissoko, G. (2019) *World Journal of Condensed Matter Physics*, **9**, 102-111. https://doi.org/10.4236/wjcmp.2019.94008

[36] Ba, M., Thiam, N., Thiame, M., Traore, Y., Diop, M., Ba, M., Sarr, C., Wade, M. and Sissoko, G. (2019) *Journal of Electromagnetic Analysis and Applications*, **11**, 173-185. https://doi.org/10.4236/jemaa.2019.1110012

[37] Kosso, A.M.M., Thiame, M., Traore, Y., Diatta, I., Ndiaye, M., Habiboullah, L., Ly, I. and Sissoko, G. (2018) *Journal of Scientific and Engineering Research*, **5**, 259-269.

Journal of Modern Physics, 2019, 10, 1606-1614
https://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

Analysis of Resonance Absorption in Multilayered Thin-Film Bi-Grating

Taikei Suyama

Department of Electrical and Computer Engineering, National Institute of Technology, Akashi College, Akashi, Hyogo, Japan
Email: suyama@akashi.ac.jp

How to cite this paper: Suyama, T. (2019) Analysis of Resonance Absorption in Multilayered Thin-Film Bi-Grating. *Journal of Modern Physics*, **10**, 1606-1614.
https://doi.org/10.4236/jmp.2019.1013106

Received: October 29, 2019
Accepted: November 24, 2019
Published: November 27, 2019

Abstract

The resonance absorption of a multilayered bi-grating which consists of thin-film corrugated periodically in two directions is investigated. The absorption in a multilayered thin-film bi-grating has been of considerable interest since we can expect more complex behaviors in the absorption phenomen by virtue of the presence of double periodicity and multilayer structure. In solving the problem, we employed a computational technique based on modal expansion. Taking a sandwiched structure /Ag/SiO_2/Ag/ for an example, we observed: 1) excitation of a single-interface surface plasmon mode at the lit surface of the 1st Ag layer with strong field enhancement for thick enough Ag layer case; 2) excitation of coupled short-range or long-range surface plasmon modes at each surface between vacuum and Ag layers with strong field enhancements for thin enough Ag layer cases no matter with the thickness of SiO_2 layers; 3) enhancements of field at surfaces between Ag and SiO_2 layers in some cases related with the thickness of SiO_2 layers. The coupled plasmon modes were resulted by the resonance waves on four surfaces in these cases.

Keywords

Multilayered Thin-Film, Surface Plasmon, Resonance Absorption, Bi-Grating

1. Introduction

Periodically corrugated thin metal films have an interesting property such as the partial or total absorption of incident light energy. The absorption is associated with the excitation of the surface plasmons and is then termed the resonance absorption [1] [2] [3]. Most of studies on the resonance absorption have mainly dealt with a thin metal film grating whose surfaces are periodic in one direction [4] [5] [6].

In our previous study, we examined the excitation of coupled plasmon modes

in a thin-film grating made of a metal [7]. When the metal is thick, e.g., more than ten times the skin depth, the plasmon can be excited on the lit surface alone. This is termed a single-interface surface plasmon (SISP). When the thickness is decreased, the plasmon can be seen also on the other surface of the film. The two plasmon waves interact with each other to form two coupled plasmon modes called short-range and long-range surface plasmon (SRSP and LRSP) [8].

In the present research, we consider a sandwiched structure: metal/dielectric/metal, which is interesting for the application in development of optical equipment, for example improving sensitivity of clinical sensing [9], surface enhanced phenomena such as Ramman scattering, and solar cells.

2. Formulation and Method of Solution

In this section we first formulate the problem of diffraction by multilayered thin-film bi-grating shown in **Figure 1(a)**. After formulating the problem, we state a method of solution based on a modal-expansion approach.

2.1. Incident Wave

The electric and magnetic field of an incident light is given by

$$\begin{bmatrix} \boldsymbol{E}^{\mathrm{i}} \\ \boldsymbol{H}^{\mathrm{i}} \end{bmatrix}(\mathrm{P}) = \begin{bmatrix} \boldsymbol{e}^{\mathrm{i}} \\ \boldsymbol{h}^{\mathrm{i}} \end{bmatrix} \exp\left(i\boldsymbol{k}^{\mathrm{i}} \cdot \boldsymbol{P}\right) \tag{1}$$

where $\boldsymbol{e}^{\mathrm{i}}$ and $\boldsymbol{h}^{\mathrm{i}}$ are the electric- and magnetic-field amplitude; $\boldsymbol{k}^{\mathrm{i}} = [\alpha, \beta, -\gamma]$ is the incident wave vector with $\alpha = n_0 k^{\mathrm{i}} \sin\theta\cos\varphi$, $\beta = n_0 k^{\mathrm{i}} \sin\theta\sin\varphi$, $\gamma = n_0 k^{\mathrm{i}} \cos\theta$, $k^{\mathrm{i}} = 2\pi/\lambda$ and n_0 is the relative refractive index of region V_0; P = (X, Y, Z) is an observation point; λ is the wavelength of the incident wave; θ is the incident angle between the Z-axis and the incident wave-vector; φ is the azimuth angle between the X-axis and the plane of incidence.

The amplitude of the incident electric field can be decomposed into TE- (TM-) component, which means the electric (or magnetic) field is perpendicular to the plane of incidence. To do this, we define two unit vectors $\boldsymbol{e}^{\mathrm{TE}}$ and $\boldsymbol{e}^{\mathrm{TM}}$ that span a plane orthogonal to $\boldsymbol{k}^{i}$. Hence, the amplitude $\boldsymbol{e}^{\mathrm{i}}$ in (1) is decomposed as

$$\boldsymbol{e}^{\mathrm{i}} = \boldsymbol{e}^{\mathrm{TE}}\cos\delta + \boldsymbol{e}^{\mathrm{TM}}\sin\delta \tag{2}$$

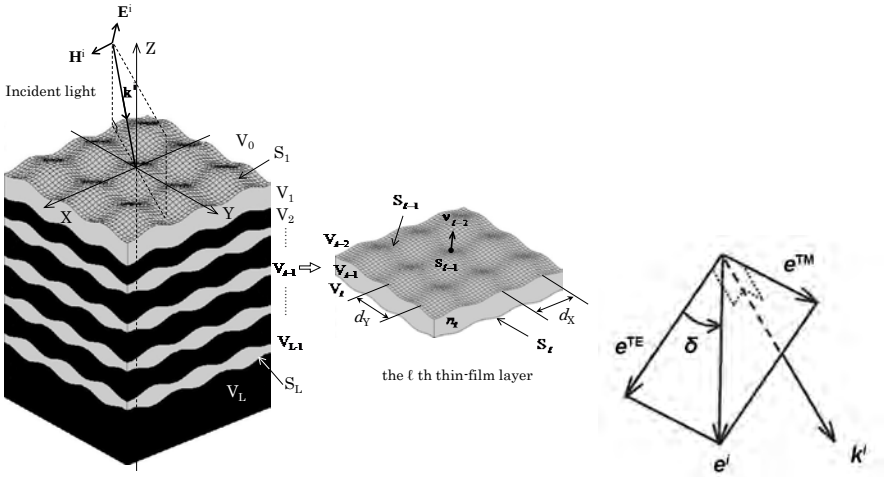
where the symbol δ is the polarization angle between $\boldsymbol{e}^{\mathrm{i}}$ and $\boldsymbol{e}^{\mathrm{TE}}$ shown in **Figure 1(b)**.

2.2. Diffracted Wave

We seek for the diffracted fields $\boldsymbol{E}_l$ (P) and $\boldsymbol{H}_l$ (P) in each region. These should satisfy the following requirements.

(C1) The Helmholtz equations in each region;

(C2) Radiation conditions: The diffracted waves in V_0 (or V_L) should propagate or attenuate in the positive (or negative) Z-direction;



(a) (b)

Figure 1. (a) Schematic representation of multilayered thin-film bi-grating with an incident light $\left(\mathrm{V}_0:\left\{\mathrm{P}\,|\,Z>\mathrm{S}_1(X,Y)\right\},\mathrm{V}_\mathrm{L}:\left\{\mathrm{P}\,|\,Z<\mathrm{S}_L(X,Y)\right\}\right)$; (b) Definition of a polarization angle.

(C3) Periodicity conditions: Each component of the diffracted electric and magnetic field should satisfy

$$f(X+d,Y,Z)=\exp(i\alpha d)f(X,Y,Z), \tag{3}$$

$$f(X,Y+d,Z)=\exp(i\beta d)f(X,Y,Z) \tag{4}$$

where, α and β are the phase constants in X and Y.

(C4) Boundary conditions: The tangential components of electric and magnetic fields are continuous across the boundaries S_l.

2.3. Geometry of the Multilayered Thin-Film Bi-Grating

The multilayered thin-film bi-grating which is laminating $L-1$ grating layers has period d in both X- and Y-directions shown in **Figure 1(a)**. The semi-infinite region over the multilayered bi-grating and the substrate are denoted by V_0 and V_L, respectively. Moreover, each region inside the multilayered bi-grating, which is numbered starting from the incident side, is denoted by $\mathrm{V}_\ell\left(\ell=1,2,\cdots,L-1\right)$. The regions $\mathrm{V}_\ell\left(\ell=0,1,\cdots,L\right)$ are filled with isotropic and homogeneous media with refractive indices n_ℓ, and a permeability of each region is equal with that of the vacuum μ_0. The interface between V_ℓ and $\mathrm{V}_{\ell+1}$ is denoted by S_ℓ $\left(\ell=1,2,\cdots,L\right)$. The profile of S_ℓ is sinusoidal and given by

$$z=\eta_\ell(x,y)=\frac{H_l}{2}\left\{\cos\left(\frac{2\pi x}{d_x}\right)+\cos\left(\frac{2\pi y}{d_y}\right)\right\}-\sum_{i=1}^{\ell-1}\omega_i \tag{5}$$

where ω_ℓ denotes an average distance between S_ℓ and S_{l+1}. The value H_ℓ is regarded as a groove depth of the boundary S_ℓ

2.4. Method of Solution

We solve the problem above using Yasuura's method of modal expansion [10]. To do this, we first define the set of modal functions; next we construct approximate solutions in terms of finite modal expansions with unknown coefficients; and, finally we determine the coefficients applying the boundary conditions.

Modal functions: Because the diffracted waves have both TE- and TM-components, we need TE and TM vector modal functions in constructing the solutions. Here we employ the functions derived from the Floquet modes (separated solutions of the Helmholtz equations satisfying the periodicity (C3) and the radiation conditions (C2) if necessary). The modal functions for electric fields for each region are given by

$$\boldsymbol{\varphi}_{\ell mn}^{\mathrm{TE,TM}\pm}(\mathrm{P}) = \boldsymbol{e}_{\ell mn}^{\mathrm{TE,TM}\pm} \exp\left(i\boldsymbol{k}^{\mathrm{i}} \cdot \boldsymbol{P}\right) \tag{6}$$

where, $m,n = 0,\pm1,\pm2,\cdots$ and $l = 0,1,2,\cdots,L$, and wave vectors in (6) are defined by

$$\boldsymbol{e}_{\ell mn}^{\mathrm{TE}\pm} = \frac{\boldsymbol{k}_{\ell mn}^{\pm} \times \boldsymbol{i}_Z}{\left|\boldsymbol{k}_{\ell mn}^{\pm} \times \boldsymbol{i}_Z\right|},\ \boldsymbol{e}_{\ell mn}^{\mathrm{TM}\pm} = \frac{\boldsymbol{e}_{\ell mn}^{\mathrm{TE}\pm} \times \boldsymbol{k}_{\ell mn}^{\pm}}{\left|\boldsymbol{e}_{\ell mn}^{\mathrm{TE}\pm} \times \boldsymbol{k}_{\ell mn}^{\pm}\right|} \tag{7}$$

and

$$\begin{aligned} &\boldsymbol{k}_{\ell mn}^{\pm} = \left[\alpha_m, \beta_n, \pm\gamma_{\ell mn}\right],\ \alpha_m = \alpha + \frac{2m\pi}{d_x},\ \beta_n = \beta + \frac{2n\pi}{d_y}, \\ &\gamma_{\ell mn} = \left(n_\ell^2 k^2 - \alpha_m^2 - \beta_n^2\right)^{1/2} \end{aligned} \tag{8}$$

where $\mathrm{Re}\left(\gamma_{\ell mn}\right) \geq 0$ and $\mathrm{Im}\left(\gamma_{\ell mn}\right) \geq 0$. We use the modal functions defined in equations from (6) to (8) to construct approximations of diffracted electric fields. For the accompanying magnetic fields, we employ

$$\boldsymbol{\psi}_{\ell mn}^{\mathrm{TE,TM}\pm}(\mathrm{P}) = \frac{1}{\omega\mu_0} \boldsymbol{k}_{\ell mn}^{\pm} \times \boldsymbol{\varphi}_{\ell mn}^{\mathrm{TE,TM}\pm} \tag{9}$$

Approximate solutions: To satisfy the radiation condition (C2), the approximate solution in V_0 should have a form of finite linear combination of up-going modal functions with unknown coefficients. Likewise, the solution in V_L must be a linear combination of down-going modal functions. The solution in V_b, however, must have both up- and down-going waves. To show the traveling direction of a modal function, we use superscripts + and − representing up- and down-going waves. Here, we form approximate solutions for the diffracted electric and magnetic fields in V_l:

$$\begin{aligned} \begin{pmatrix} \boldsymbol{E}_{\ell N}^{\mathrm{d}} \\ \boldsymbol{H}_{\ell N}^{\mathrm{d}} \end{pmatrix}(\mathrm{P}) &= \sum_{m,n=-N}^{N} A_{\ell mn}^{\mathrm{TE}+}(N) \begin{pmatrix} \phi_{\ell mn}^{\mathrm{TE}+} \\ \psi_{\ell mn}^{\mathrm{TE}+} \end{pmatrix}(\mathrm{P}) + \sum_{m,n=-N}^{N} A_{\ell mn}^{\mathrm{TM}+}(N) \begin{pmatrix} \phi_{\ell mn}^{\mathrm{TM}+} \\ \psi_{\ell mn}^{\mathrm{TM}+} \end{pmatrix}(\mathrm{P}) \\ &+ \sum_{m,n=-N}^{N} A_{\ell mn}^{\mathrm{TE}-}(N) \begin{pmatrix} \phi_{\ell mn}^{\mathrm{TE}-} \\ \psi_{\ell mn}^{\mathrm{TE}-} \end{pmatrix}(\mathrm{P}) + \sum_{m,n=-N}^{N} A_{\ell mn}^{\mathrm{TM}-}(N) \begin{pmatrix} \phi_{\ell mn}^{\mathrm{TM}-} \\ \psi_{\ell mn}^{\mathrm{TM}-} \end{pmatrix}(\mathrm{P}), \left(\ell = 1,2,\cdots,L\right) \end{aligned} \tag{10}$$

where N denotes the truncation number.

Boundary matching: Because the approximate solutions satisfy the require-

ments (C1), (C2), and (C3) by definition, the unknown coefficients $A_{\ell mn}^{\mathrm{TE}\pm}(N)$ and $A_{\ell mn}^{\mathrm{TM}\pm}(N)$ are determined such that the solutions satisfy the boundary conditions (C4) in an approximate sense. In the Yasuura's method, the least-squares method is employed to fit the solution to the boundary conditions [10] [11]. That is, we find the coefficients that minimize the weighted mean-square error by

$$I_N = \int_{S_1'} \left| \boldsymbol{\nu} \times \left[\boldsymbol{E}_{1N}^{\mathrm{d}} + \boldsymbol{E}^{\mathrm{i}} - \boldsymbol{E}_{2N}^{\mathrm{d}} \right](s_1) \right|^2 \mathrm{d}s + |\Gamma_1|^2 \int_{S_1'} \left| \boldsymbol{\nu} \times \left[\boldsymbol{H}_{1N}^{\mathrm{d}} + \boldsymbol{H}^{\mathrm{i}} - \boldsymbol{H}_{2N}^{\mathrm{d}} \right](s_1) \right|^2 \mathrm{d}s + \sum_{l=2}^{L-1} \left\{ \int_{S_l'} \left| \boldsymbol{\nu} \times \left[\boldsymbol{E}_{lN}^{\mathrm{d}} - \boldsymbol{E}_{l+1N}^{\mathrm{d}} \right](s_l) \right|^2 \mathrm{d}s + |\Gamma_l|^2 \int_{S_l'} \left| \boldsymbol{\nu} \times \left[\boldsymbol{H}_{lN}^{\mathrm{d}} - \boldsymbol{H}_{l+1N}^{\mathrm{d}} \right](s_l) \right|^2 \mathrm{d}s \right\} \tag{11}$$

where S_l' denotes one-period cells of the interface S_ℓ, Γ_ℓ is the intrinsic impedance of the medium in V_ℓ and $\boldsymbol{\nu}$ is a unit normal vector of each boundary.

To solve the least-squares problem on a computer, we need a discretized form of the problem. We first discretize the weighted mean-square error I_N by applying a two-dimensional trapezoidal rule where the number of sampling points is chosen as $2(2N + 1)$ [9]. We then employ orthogonal decomposition methods [singular-value decomposition (SVD) and QR decomposition (QRD)] in solving the discretized problem [12] [13].

It is known that the solutions obtained by Yasuura's method have proof of convergence [13] [14]. We, therefore, can employ the coefficients $A_{\ell mn}^{\mathrm{TE}\pm}(N)$ and $A_{\ell mn}^{\mathrm{TM}\pm}(N)$ with sufficiently large N for which the coefficients are stable in evaluating diffracted fields. The power reflection and transmission coefficient of the (m, n) order propagating mode in V_0 and V_L are given by

$$R_{mn} = \frac{\gamma_{0mn}}{\gamma} \left| A_{1mn}^{\mathrm{TE}+} \right|^2 + \frac{\gamma_{0mn}}{\gamma} \left| A_{1mn}^{\mathrm{TM}+} \right|^2, \mathrm{Re}(\gamma_{0mn}) \geq 0, \tag{12}$$

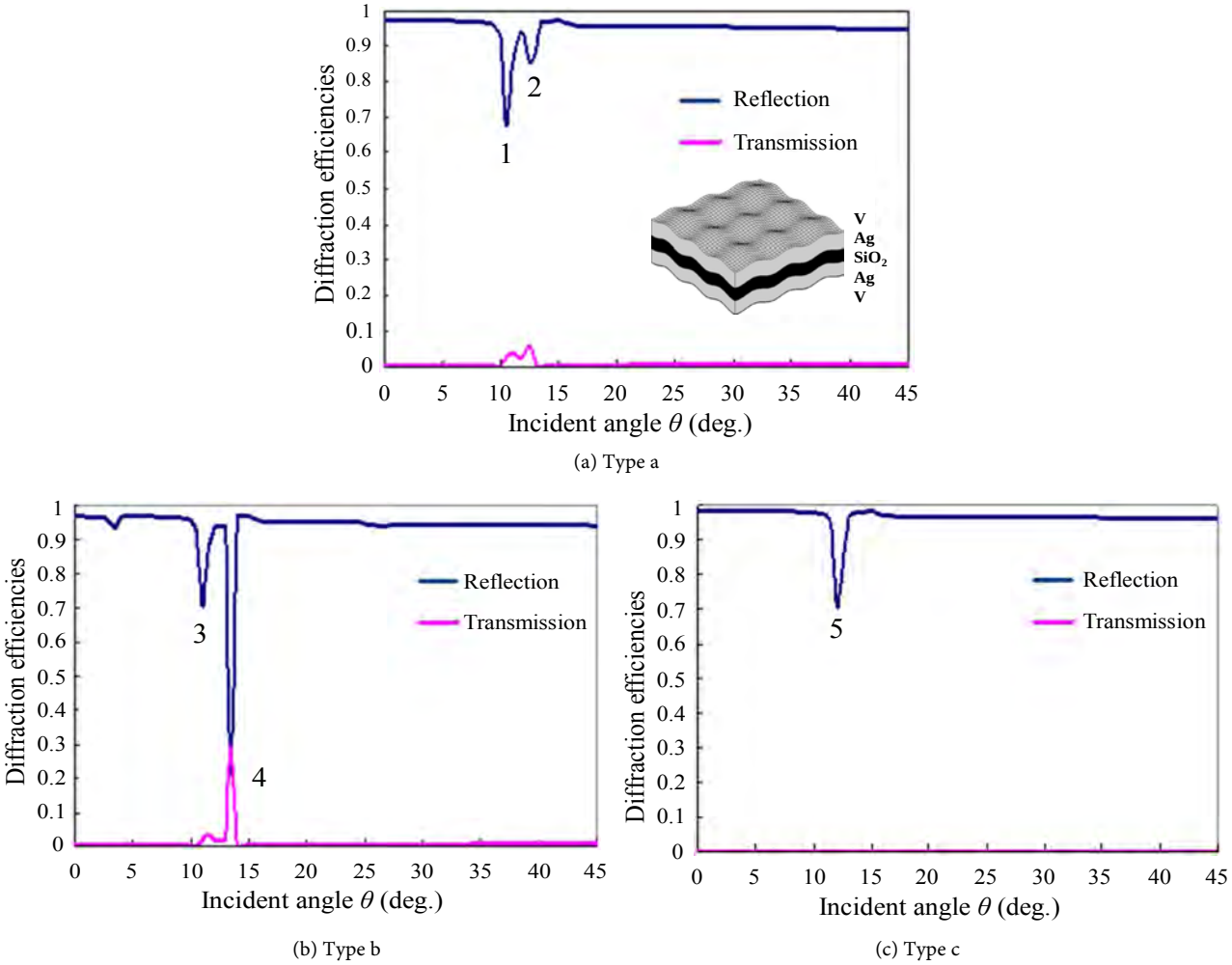
$$T_{mn} = \frac{\gamma_{Lmn}}{\gamma} \left| A_{Lmn}^{\mathrm{TE}-} \right|^2 + \frac{\gamma_{Lmn}}{\gamma} \left| A_{Lmn}^{\mathrm{TM}-} \right|^2, \mathrm{Re}(\gamma_{Lmn}) \geq 0. \tag{13}$$

The coefficient defined above is the power carried away by propagating diffraction orders normalized by the incident power.

3. Numerical Results

The multilayered bi-grating is made by 3 layers: /Ag/SiO_2/Ag/. The incident light is a TM-polarized plane wave with a 650 nm wavelength. The relative refractive index of Vacuum = 1, $n_{\mathrm{Ag}} = 0.07+4.2i$ and $n_{\mathrm{SiO_2}} = 1.5$. The periods of two directions $d_x = d_y = 556$ nm. We consider 3 types of gratings with different thickness pairs of each region: (A) $e_{\mathrm{Ag}} = e_{\mathrm{SiO_2}} = 27.8$ nm; (B) $e_{\mathrm{Ag}} = 27.8$ nm, $e_{\mathrm{SiO_2}} = 278$ nm; and (C) $e_{\mathrm{Ag}} = 278$ nm, $e_{\mathrm{SiO_2}} = 27.8$ nm. We will then calculate the diffraction efficiencies and field distributions of these gratings.

Figure 2 shows the (0, 0)-th order reflection and transmission coefficient as functions of the incident angle θ for 3 types of gratings with the azimuth angle $\phi = 45^\circ$. Five dips are observed on reflection curves throughout **Figures 2(a)-(c)**, the field distributions of the total electric fields in the vicinity of the SiO_2 layer for these dips are shown in **Figures 3(a)-(e)**. Distances in the Z direction are normalized by the wavelength.



(a) Type a

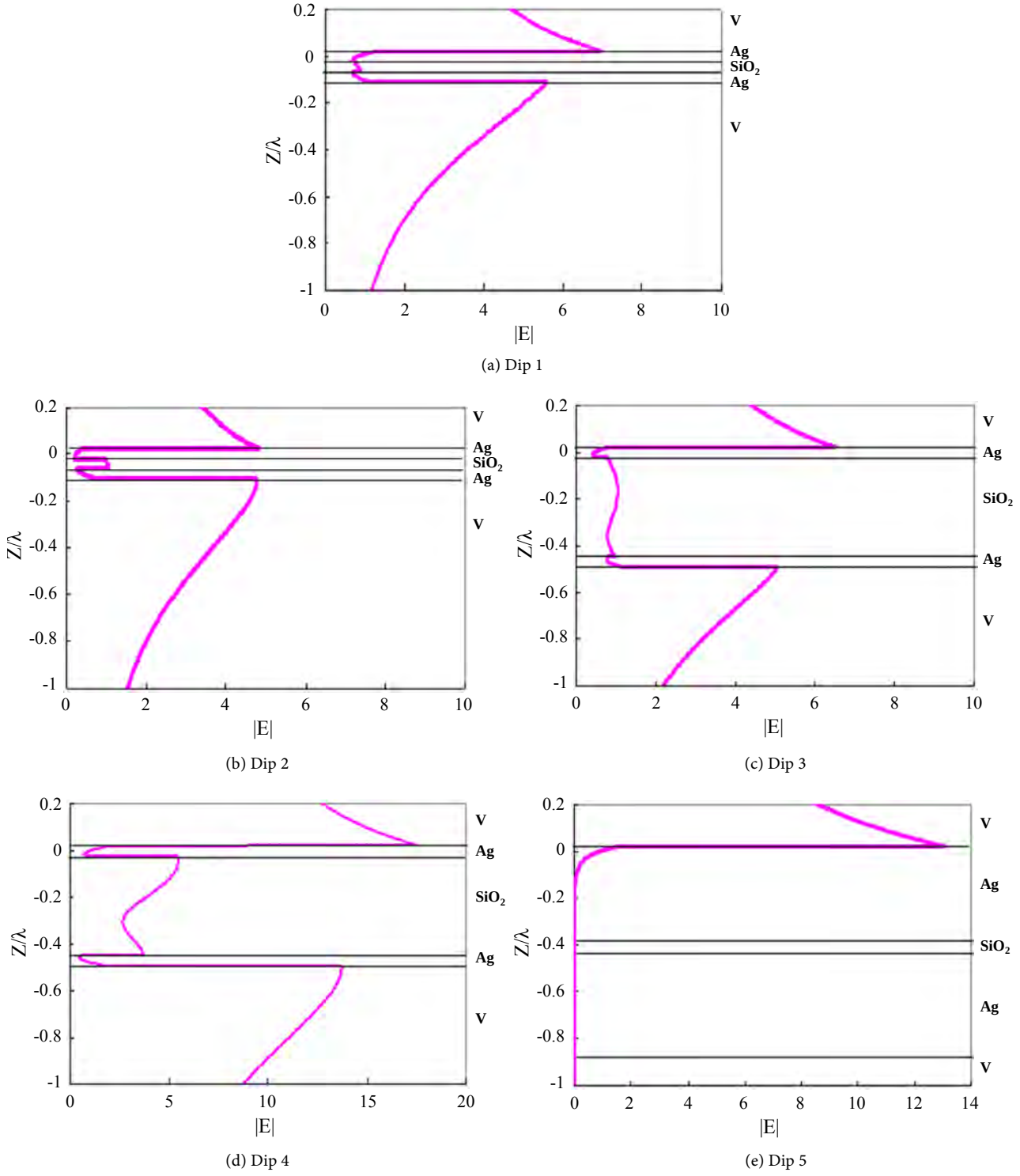
(b) Type b

(c) Type c

Figure 2. The (0, 0)-th order reflection and transmission coefficient for 3-layers doubly periodic gratings with 3 types of thickness pairs.

For type A shown in **Figure 2(a)**, two dips (dip 1 at $\theta = 106^\circ$ and dip 2 at $\theta = 12.5^\circ$) are observed on reflection curve and the transmission coefficient also increases at the same time. **Figures 3(a)** and **Figures 3(b)** show the field distributions for dips1 and 2. Strong field enhancements are observed at each surface between Vacuum and Ag layers for either of the two dips. The interaction of SPRs excited on these surfaces result in the coupled plasmon modes (SRSP or LRSP).

For type B shown in **Figure 2(b)**, two dips (dip 3 at $\theta = 11.0^\circ$ and dip 4 at $\theta = 13.5^\circ$) are observed on reflection curve and the transmission coefficient increases at the same time. Reflection at dip 4 is much lower than that of dips 1 and 2 accompanying change of the thickness of SiO_2. **Figure 3(c)** and **Figure 3(d)** show the field distributions for dips 3 and 4. Strong field enhancements are observed at each surface between Vacuum and Ag layers, similar to that of type A. In addition, fields enhance strongly at two surfaces between Ag and SiO_2 for dip 4 at the same time. This means coupled plasmon modes are resulted by resonance waves excited at four surfaces.



(a) Dip 1

(b) Dip 2

(c) Dip 3

(d) Dip 4

(e) Dip 5

Figure 3. Field distributions correspond to five dips.

For type C shown in Figure 2(c), one dip (dip 5 at $\theta = 12.0^{\circ}$) is observed on reflection curve and the transmission coefficient keeps to zero. Figure 3(e) shows the field distribution for dip 5. Field enhanced only on the lit surface and quickly attenuates through the 1st Ag layer as shown in Figure 3(e). Because the

1st Ag layer is enough thick, the oscillation near the upper surface does not reach the lower surface and, hence, the field below the 1st Ag layer is zero. This indicates the excitation of SISP.

4. Conclusion

We solved the problems for 3 thin-films bi-grating. By calculating the diffraction efficiency and field distributions, we showed that the SPR phenomenon excited and we observed: 1) excitation of a SISP mode at the lit surface of the 1st Ag layer with strong field enhancement for thick enough layer case; 2) excitation of coupled SPR modes (SRSP or LRSP) at each surface between vacuum and Ag layers with strong field enhancements for thin enough Ag layer cases no matter with the thickness of SiO_2 layers; 3) enhancements of field at surfaces between Ag and SiO_2 layers in some cases related with the thickness of SiO_2 layers. The coupled plasmon modes were resulted by the resonance waves excited on four surfaces in these cases. In future, we plan to study applications for multilayered bi-grating such as improving the sensitivity of a bio-sensor by determining changes of SPRs excited at different layers' surfaces.

Conflicts of Interest

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

References

[1] Raether, H. (1977) *Physics of Thin Films*, **9**, 145-261.

[2] Raeter, H. (1982) Surface Plasmon and Roughness. In: Agranovich, V.M. and Mills, D.L., Eds., *Surface Polaritons*, North-Holland, Amsterdam, 331-403. https://doi.org/10.1016/B978-0-444-86165-8.50015-3

[3] Nevier, M. (1980) The Homogenous Problem. In: Petited, R., Ed., *Electromagnetic Theory of Gratings*, Springer-Verlag, Berlin, 123-157. https://doi.org/10.1007/978-3-642-81500-3_5

[4] Economou, E.N. (1969) *Physical Review*, **182**, 539-554. https://doi.org/10.1103/PhysRev.182.539

[5] Chen, Z., Hooper, I.R. and Sambles, J.R. (2008) *Physical Review B*, **77**, Article ID: 161405. https://doi.org/10.1103/PhysRevB.77.161405

[6] Maksimenko, L., Matyash, I., Mischuk, O., Rudenko, S., Serdega, B. and Stetsenko, M. (2014) *Journal of Modern Physics*, **5**, 617-626. https://doi.org/10.4236/jmp.2014.58072

[7] Suyama, T., Zhang, Y.J., Okuno, Y., Luo, Z.Q. and Matsuda, T. (2010) *PIERS Online*, **6**, 76-80. https://doi.org/10.2529/PIERS090902043932

[8] Hibbins, A.P., Murray, W.A., Tyler, J., Wedge, S., Barnes, W.L. and Sambles, J.R. (2006) *Physical Review B*, **74**, Article ID: 073408. https://doi.org/10.1103/PhysRevB.74.073408

[9] Ahn, J.H., Seong, T.Y., Kim, W.M., Lee, T.S., Kim, I. and Lee, K.S. (2012) *Optics Express*, **20**, 21729-21738. https://doi.org/10.1364/OE.20.021729

[10] Yasuura, K. and Itakura, T. (1966) Approximation Method for Wave Functions (I),

(II), and (III). Kyushu Univ. Tech. Rep., Vol. 38, 72-77, 1965; Vol. 38, 378-385, 1966; Vol. 39, 51-56.

[11] Okuno, Y. (1990) Mode-Matching Method. In: Yamashita, E., Ed., *Analysis Methods for Electromagnetic Wave Problems*, Artech House, Boston, 107-138.

[12] Lawson, C.L. and Hanson, R.J. (1974) Solving Least Squares Problem. Prentice-Hall, Englewood Cliffs.

[13] Suyama, T., Okuno, Y., Matsushima, A. and Ohtsu, M. (2008) *Progress in Electromagnetics Research B*, **2**, 83-102. https://doi.org/10.2528/PIERB07110301

[14] Matsuda, T. and Okuno, Y. (1996) *Radio Science*, **31**, 1791-1798. https://doi.org/10.1029/96RS02153

Call for Papers

Journal of Modern Physics

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

Journal of Modern Physics **(JMP)** is an international journal dedicated to the latest advancement of modern physics. The goal of this journal is to provide a platform for scientists and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of modern physics.

Editor-in-Chief

Prof. Yang-Hui He City University, UK

Subject Coverage

Journal of Modern Physics publishes original papers including but not limited to the following fields:

Biophysics and Medical Physics
Complex Systems Physics
Computational Physics
Condensed Matter Physics
Cosmology and Early Universe
Earth and Planetary Sciences
General Relativity
High Energy Astrophysics
High Energy/Accelerator Physics
Instrumentation and Measurement
Interdisciplinary Physics
Materials Sciences and Technology
Mathematical Physics
Mechanical Response of Solids and Structures
New Materials: Micro and Nano-Mechanics and Homogeneization
Non-Equilibrium Thermodynamics and Statistical Mechanics
Nuclear Science and Engineering
Optics
Physics of Nanostructures
Plasma Physics
Quantum Mechanical Developments
Quantum Theory
Relativistic Astrophysics
String Theory
Superconducting Physics
Theoretical High Energy Physics
Thermology

We are also interested in: 1) Short Reports—2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data; 2) Book Reviews—Comments and critiques.

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

https://www.scirp.org/journal/jmp **E-mail: jmp@scirp.org**

What is SCIRP?

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

What is Open Access?

All original research papers published by SCIRP are made freely and permanently accessible online immediately upon publication. To be able to provide open access journals, SCIRP defrays operation costs from authors and subscription charges only for its printed version. Open access publishing allows an immediate, worldwide, barrier-free, open access to the full text of research papers, which is in the best interests of the scientific community.

- High visibility for maximum global exposure with open access publishing model
- Rigorous peer review of research papers
- Prompt faster publication with less cost
- Guaranteed targeted, multidisciplinary audience

Scientific Research Publishing

Website: https://www.scirp.org

Subscription: sub@scirp.org

Advertisement: service@scirp.org