ISSN: 2327-4352 Volume 7, Number 8, August 2019

Scientific Research Publishing

Journal of Applied Mathematics and Physics

ISSN: 2327-4352

www.scirp.org/journal/jamp

JOURNAL EDITORIAL BOARD

ISSN 2327-4352 (Print) ISSN 2327-4379 (Online)
http://www.scirp.org/journal/jamp

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1687-1910
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Table of Contents

Volume 7 Number 8 August 2019

Journal of Applied Mathematics and Physics, 2019, 7, 1687-1910
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Journal of Applied Mathematics and Physics (JAMP)
Journal Information

SUBSCRIPTIONS

The *Journal of Applied Mathematics and Physics* (Online at Scientific Research Publishing, www.SciRP.org) is published monthly by Scientific Research Publishing, Inc., USA.

Subscription rates:
Print: $39 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: jamp@scirp.org

Journal of Applied Mathematics and Physics, 2019, 7, 1687-1696
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

DNA Sequences with Forbidden Words and the Generalized Cantor Set

Zhuowei Yang[1], Ping Wang[2]

[1]College of Mathematics and Statistics, Yunnan University, Kunming, China
[2]Department of Mathematics, Penn State University, Schuylkill Haven, PA, USA
Email: yangzhuowei@outlook.com, pxw10@psu.edu

How to cite this paper: Yang, Z.W. and Wang, P. (2019) DNA Sequences with Forbidden Words and the Generalized Cantor Set. *Journal of Applied Mathematics and Physics*, **7**, 1687-1696.
https://doi.org/10.4236/jamp.2019.78115

Received: June 28, 2019
Accepted: August 9, 2019
Published: August 12, 2019

Abstract

In this work, we establish relations between DNA sequences with missing subsequences (the forbidden words) and the generalized Cantor sets. Various examples associated with some generalized Cantor sets, including Hao's frame representation and the generalized Sierpinski Set, along with their fractal graphs, are also presented in this work.

Keywords

DNA Sequences, Generalized Cantor Set

1. Introduction

Researchers have been interested in the relationships between fractals and DNA structures for years. Just recently, Anitas and Slyamov [1] studied multiscale fractal representing DNA sequences using small-angle scattering analysis. Cattani and Pierro [2] conducted a multifractal analysis of binary images of DNA in order to define a methodological approach to the classification of DNA sequences. Badea and her collaborators [3] characterized the geometry of some medical images of tissues in terms of complexity parameters such as the fractal dimension (FD). Carlo Cattani presented analysis of DNA based on the indicator matrix together with some elementary approach to a fractal estimate of DNA sequences in the book [4] edited by Elloumi and Zomaya. Albrecht-Buehler [5] identified explicitly the GA-sequences as a class of fractal genomic sequences. Ainsworth [6] investigated how the cell's nucleus holds molecules that manage human's DNA in the right location. In a book edited by Crilly, Earnshaw and Jones, Voss applied standard spectral density measurement techniques to demonstrate the ubiquity of low frequency noise and long range fractal correlations.

The study of the genome or DNA sequences through fractal analysis is very interesting. DNA sequences can be seen as sequences over the alphabet $\Sigma = \{a, c, g, t\}$. Subsequences that do not appear in DNA are considered as forbidden words. A visualization method of the forbidden words in [7] [8] [9] [10] [11] has been designed by B.-L. Hao since 2000. This method is now called Hao's frame representation. Recently, C.-X. Huang and S.-L. Peng discussed this method in detail, and many beautiful graphics were provided in [12] [13]. From these geometric intuitions, it can be observed that these forbidden words demonstrate certain fractal properties. In fact in this work we generated some amazing fractal graphs associated with DNA sequences with forbidden words as shown in **Figure 1**.

It is important to explore the fractal generating mechanism that is associated with the forbidden words in the sequence. H. J. Jeffrey [14] [15] and P. Tiňo [16] [17] tried to associate the forbidden words with the IFS (Iterated Functions Systems) using chaos game algorithm. Denote Σ^* as the set of all finite sequences over Σ. Then how to find a generating formula or the mapping $\sigma : \Sigma \to w \in \Sigma^*$, where w is a sequence that does not contain forbidden subsequences, or corresponding iteration method? As was pointed out by P. Tiňo, the IFS is a multifractal and therefore the generating formula would be relatively complicated.

In order to detect the structures of some symbolic sequences, one has to find the properties of their topology and metric and be able to visualize these sequences. To do this, we have to provide a type of graphical representation together with their topology and metric properties so that we can directly reveal their corresponding fractal graphs. This kind of representation method is important and necessary.

For an alphabet with cardinal 3, the well known CGR method (that is, Chaos Game Representation method) was first introduced by M.F. Barnsley by considering the points in an equilateral triangle. The substrings of a string were shown graphically (see [18]). For an alphabet $\Sigma = \{a, c, g, t\}$ with cardinality 4, the CGR method was later generalized by H.J. Jeffrey so that the DNA sequences can be visualized (see [14] [15]). The authors have transformed the DNA sequences into pseudo random walk in a 2-dimensional plane or in a 3-dimensional space [19] [20] [21]. We notice here that an iterated function system can be applied to construct a graphical representation of some DNA sequences [16] [17]. The points in the unit square $[0,1]\times[0,1]$ can be used to denote the substrings of the DNA sequences. Consequently, the four vertices of the unit square are labelled as a, c, g, t.

In application, the frame representation method proposed by Hao *et al.* is more intuitive and visual [9] [10]. The unit square $[0,1]\times[0,1]$ is divided equally with vertical and horizontal lines so that there are 4^k congruent small squares with side length 2^{-k} and area 4^{-k}. For the alphabet $\Sigma = \{a, c, g, t\}$ with cardinality 4, each small square of side length 2^{-k} is used to denote the string in Σ^k $(k = 1, 2, 3, \cdots)$ regularly (See 1-, 2- and 3-frame graphs in **Figures 2(a)-(c)**).

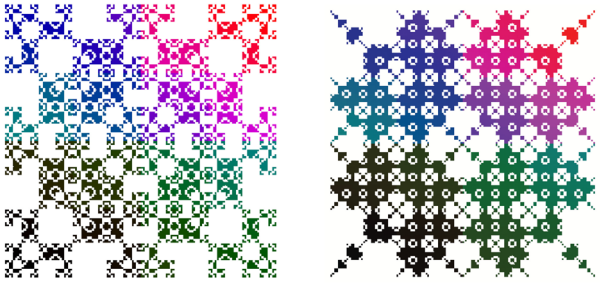
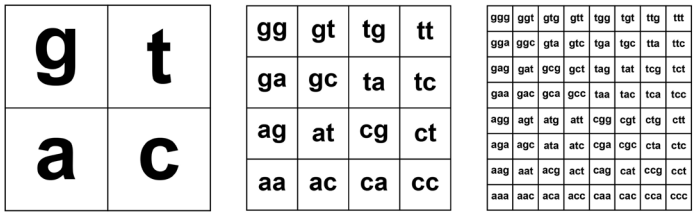


Figure 1. Graphs of some forbidden words.



(a) (b) (c)

Figure 2. The frame representation method of B.L. Hao *et al.* (a) 1-frame graph; (b) 2-frame graph; (c) 3-frame graph.

With the frame representation method of B.L. Hao, the repetition topology structure of the subsequences (*i.e.* the strings in Σ^k) of a DNA sequence can be easily visualized and efficiently drawn. The avoided or the under-represented short strings in the genome sequence form the forbidden words. These forbidden words are the reasons or the basis of the constructed fractals.

P. Tino [16] [17] proved the equivalence of the CGR method and the frame representation method of B.L. Hao *et al.* He noted that the cardinality of an alphabet can be generalized to a square integer ($|\Sigma| = b^2$ simultaneously for some integer b). We will in this paper extend the above methods and relax the restriction to the cardinality of an alphabet.

The order of this paper is as follows. In Section 2, we will first convert the problem into the discussion on certain type of generalized Cantor set, which can naturally correspond to multifractals, and then in Section 3, we will induce Hao's frame representation according to the principle that the correspondence between line segment and unit square is one-to-one [22]. Several examples, along with their fractal graphs, of some generalized Cantor sets are given at the end of this paper.

2. Forbidden Words and the Generalized Cantor Set

Rewrite the alphabet as $\Sigma = \{0,1,2,3\}$. We first give the following definition.

Definition 2.1 *Let* $\Sigma = \{0,1,2,3\}$. *Denote B as the set consist of l finite sequences with length* $k(\geq 1)$:

$$B = \{t_{11}\cdots t_{1k}, t_{21}\cdots t_{2k}, \cdots, t_{l1}\cdots t_{lk}\}, \quad t_{ij} \in \Sigma, i = 1,\cdots,l, i = 1,\cdots,k. \tag{1}$$

Then call the infinite sequences over Σ

$$s = a_1 a_2 \cdots a_n \cdots,\ a_n \in \Sigma,\ \text{and}\ a_1 a_{i+1} \cdots a_{i+k-1} \notin B,\ i = 1,2\cdots \tag{2}$$

the DNA sequence with no forbidden words B, a.k.a. allowed sequence.

It is known that when $x \in [0,1]$ is expanded in ternary representation, the subset in $[0,1]$

$$C = \left\{ x = 0_3.x_1 x_2 \cdots x_n \cdots, x \in \{0,2\} \right\}$$

is called the Cantor set. Similarly, with quaternary expansion, we give the following definition.

Definition 2.2 *When* $x \in [0,1]$ *is represented in quaternary expansion*

$$x = \sum_{n=1}^{\infty} \frac{a_n}{4^n} = 0_4.a_1 a_2 \cdots a_n \cdots,\ a_n \in \Sigma, \tag{3}$$

we call

$$C_G = \left\{ x = 0_4.a_1 a_2 \cdots a_n \cdots, a_n \in \Sigma, a_i a_{i+1} \cdots a_{i+k-1} \notin B, i = 1,2,\cdots \right\},\ k \geq 1 \tag{4}$$

the generalized Cantor set.

Apparently, the discussions on DNA sequences (1) (2) that contain no forbidden words B can be converted into the discussion on the generalized Cantor set C_G.

Let $b_k^i = 4^{k-1} t_{i1} + \cdots + 4 t_{ik-1} + t_{ik}$, $b_k^i \in \left\{ 0,1,\cdots,4^k - 1 \right\}$, $i = 1,2,\cdots,l$, and

$$B' = \left\{ b_k^1, b_k^2, \cdots, b_k^l \right\}, \tag{5}$$

Then, the condition $a_i a_{i+1} \cdots a_{i+k-1} \notin B, i = 1,2,\cdots$ in Definition 2.2 can be rewritten as

$$4^{k-1} a_i + 4^{k-2} a_{i+1} + \cdots + a_{i+k} \notin B',\ i = 1,2,\cdots$$

Theorem 2.1 *The generalized Cantor set* C_G *can be inducted by using an iteration method.*

Proof. In fact, for the $(k-1)$ th step of the quaternary expansion of $x \in [0,1]$, there is

$$x = \frac{a_1}{4} + \cdots + \frac{a_{k-1}}{4^{k-1}} + \frac{x_{k-1}}{4^{k-1}},\ 0 \leq x_{k-1} \leq 1,\ a_1, \cdots, a_k \in \Sigma. \tag{6}$$

Let

$$x_{k-1} = \frac{a_k}{4} + \frac{x_k}{4},\ 0 \leq x_k \leq 1, a_k \in \Sigma\ \text{and}\ 4^{k-1} a_1 + \cdots + a_k \notin B'. \tag{7}$$

Substitute (7) into (6),

$$x = \frac{a_1}{4} + \cdots + \frac{a_k}{4^k} + \frac{x_k}{4^k},\ 0 \leq x_k \leq 1, 4^{k-1} a_1 + \cdots + a_k \notin B'. \tag{8}$$

In general, we let

$$x_{i-1+k} = \frac{a_{i+k}}{4} + \frac{x_{i+k}}{4}, 0 \leq x_{i+k} \leq 1, a_{i+k} \in \Sigma\ \text{and}\ 4^{k-1} a_i + \cdots + a_{k+i} \notin B'. \tag{9}$$

and as $i \to \infty$, we obtain the generalized Cantor set C_G (2.2).

$\left[\sum_{j=1}^{k} \frac{a_j}{4^j}, \sum_{j=1}^{k} \frac{a_j}{4^j} + \frac{1}{4^k} \right]$, $a_j \in \Sigma$ are 4^k intervals in $[0,1]$ with length $\frac{1}{4^k}$.

From the iteration Equation (7) in the theorem, the iteration acts differently on the l subintervals than on the $4^k - l$ intervals. Hence we have [11].

Corollary 2.3 *The generalized Cantor set* C_G *is multifractal.*

Proof. In the construction of the generalized Cantor sets C_G, measures on removed portions are redistributed to the neighboring sections repeatedly. Thus C_G is multifractal.

Obviously, the generalized Cantor sets are applicable for all p-carry representation (p is an integer).

3. The Hao's Frame Representation of the Generalized Cantor Set C_G

The theoretic foundation of the construction of DNA sequences can be seen in [12]. The subintervals in the quaternary expansion of $x \in [0,1]$ can be one-to-one corresponding to the subsquares that are obtained by repeatedly equally dividing the unit square (and its subsquares) into 4 smaller subsquares. Cantor sets are created in one dimension in $[0,1]$ while Sierpinski sets are constructed in two dimension within $[0,1]\times[0,1]$. Using the corresponding relationship between the unit interval and the unit square, we can convert the discussion on the generalized Cantor sets into the discussion on the generalized Sierpinski sets on the unit square.

Let $0 \le \xi, \eta \le 1$. The binary expansion of (ξ, η) is

$$(\xi,\eta) = \left(\sum_{n=1}^{\infty}\frac{c_n}{2^n}, \sum_{n=1}^{\infty}\frac{d_n}{2^n}\right), c_n, d_n \in \{0,1\}. \tag{10}$$

The expansion can be related to the quaternary expansion of $x \in [0,1]$ as follows:

$$a_i = c_i + 2d_i = \begin{cases} 0 & :(c_i,d_i) = (0,0) \\ 1 & :(c_i,d_i) = (1,0) \\ 2 & :(c_i,d_i) = (0,1) \\ 3 & :(c_i,d_i) = (1,1) \end{cases} \tag{11}$$

Thus the forbidden words b_k^i in B' can be represented as

$$\begin{aligned} &b_k^i = 4^{k-1}\left(c_{i_1} + 2d_{i_1}\right) + \cdots + 4\left(c_{i_{k-1}} + 2d_{i_{k-1}}\right) + \left(c_{i_k} + 2d_{i_k}\right), \\ &c_{i_1}, d_{i_1}, \cdots, c_{i_k}, d_{i_k} \in \{0,1\}, i = 1,2,\cdots,l \end{aligned} \tag{12}$$

Definition 3.1 *Let* $0 \le \xi, \eta \le 1$ *and the binary expansion of* (ξ, η) *is* (10). *Then call*

$$\begin{aligned} S_G = \Bigg\{ &(\xi,\eta) = \left(\sum_{n=1}^{\infty}\frac{c_n}{2^n}, \sum_{n=1}^{\infty}\frac{d_n}{2^n}\right), c_n, d_n \in \{0,1\}, \\ &4^{k-1}\left(c_{i_1} + 2d_{i_1}\right) + \cdots + 4\left(c_{i_{k-1}} + 2d_{i_{k-1}}\right) + \left(c_{i_k} + 2d_{i_k}\right) \notin B' \Bigg\} \end{aligned} \tag{13}$$

the generalized Sierpinski set that corresponds to the the generalized Cantor set $C_G(4)$.

Theorem 3.1 *The generalized Sierpinski set* S_G *can be inducted by iterating method.*

Proof. The $(k-1)$ th binary expansion of (ξ,η) is

$$(\xi,\eta)=\left(\sum_{n=1}^{k-1}\frac{c_n}{2^n}+\frac{\xi_{k-1}}{2^{k-1}},\sum_{n=1}^{k-1}\frac{d_n}{2^n}+\frac{\eta_{k-1}}{2^{k-1}}\right),\ 0\le\xi_{k-1},\eta_{k-1}\le 1. \tag{14}$$

Let

$$\begin{aligned}&(\xi_{k-1},\eta_{k-1})=\left(\frac{c_k}{2}+\frac{\xi_k}{2},\frac{d_k}{2}+\frac{\eta_k}{2}\right),c_k,d_k\in\{0,1\},\ 0\le\xi_k,\eta_k\le 1,\\&4^{k-1}(c_1+2d_1)+\cdots+4(c_{k-1}+2d_{k-1})+(c_k+2d_k)\notin B',\ c_k,d_k\in\{0,1\}\end{aligned} \tag{15}$$

Substitute (15) into (14), we have

$$\begin{aligned}&(\xi,\eta)=\left(\sum_{n=1}^{k}\frac{c_n}{2^n}+\frac{\xi_k}{2^k},\sum_{n=1}^{k}\frac{d_n}{2^n}+\frac{\eta_k}{2^k}\right),\\&4^{k-1}(c_1+2d_1)+\cdots+4(c_{k-1}+2d_{k-1})+(c_k+2d_k)\notin B'\end{aligned}$$

Generally, let

$$\begin{aligned}&(\xi_{k+i-1},\eta_{k+i-1})=\left(\frac{c_{k+i}}{2}+\frac{\xi_{k+i}}{2},\frac{d_{k+i}}{2}+\frac{\eta_{k+i}}{2}\right),\ 0\le\xi_{k+i},\eta_{k+i}\le 1,\ c_{k+i},d_{k+i}\in\{0,1\},\\&4^{k-1}(c_i+2d_i)+\cdots+4(c_{k+i-1}+2d_{k+i-1})+(c_{k+i}+2d_{k+i})\notin B'\end{aligned}$$

Noticing the corresponding relationship between numbers and the subsquares, naturally we have Hao's frame representation. The second-order Hao's frame representation can be inducted from the corresponding relationship illustrated in Figure 3.

The next few examples illustrate analytic structure of some DNA sequences along with the fractal graphs of the relevant generalized Cantor sets.

Example 3.2 *Let* $\Sigma=\{0,1,2,3\}$, $B=\{00,11,22\}$. *Then* $B'=\{0,5,10\}$. *Hence the arithmetic expression of the generalized Cantor set is*

$$x=0_4.a_1a_2\cdots a_n\cdots,\ a_n\in\{0,1,2,3\}\ \text{and}\ 4a_i+a_{i+1}\ne 0,5,10,\ i=1,2,\cdots$$

And the symbolic sequence is

$$a_1a_2\cdots a_n\cdots,\ \ a_n\in\Sigma,\ a_ia_{i+1}\notin B,\ i=1,2,\cdots$$

which is shown graphically in Figure 4.

Example 3.3 *Let* $\Sigma=\{0,1,2,3\}$, $B=\{10,20,30\}$. *Then* $B'=\{4,8,12\}$. *Hence the arithmetic expression of the generalized Cantor set is*

$$x=0_4.a_1a_2\cdots a_n\cdots,\ a_n\in\{0,1,2,3\}\ \text{and}\ 4a_i+a_{i+1}\ne 4,8,12,\ i=1,2,\cdots$$

And the symbolic sequence is

$$a_1a_2\cdots a_n\cdots,\ \ a_n\in\Sigma,\ a_ia_{i+1}\notin B,\ i=1,2,\cdots$$

with graphs Figure 5:

Example 3.4 *Let* $\Sigma=\{0,1,2,3\}$, $B=\{011,022,100,133,200,233,311,322\}$. *Then* $B'=\{5,10,16,31,32,47,53,58\}$. *Hence the arithmetic expression of the generalized Cantor set is*

$$x=0_4.a_1a_2\cdots a_n\cdots,\ a_n\in\{0,1,2,3\}\ \text{and}\ 4^2a_i+4a_{i+1}+a_i\notin B',\ i=1,2,\cdots$$

And the symbolic sequence is

$$a_1 a_2 \cdots a_n \cdots, \quad a_n \in \Sigma, a_i a_{i+1} \notin B, i = 1, 2, \cdots$$

which are shown below

Similarly, we could produce the following amazing fractal graphs shown in Figure 6, Figure 7, of different DNA sequences with various forbidden words.

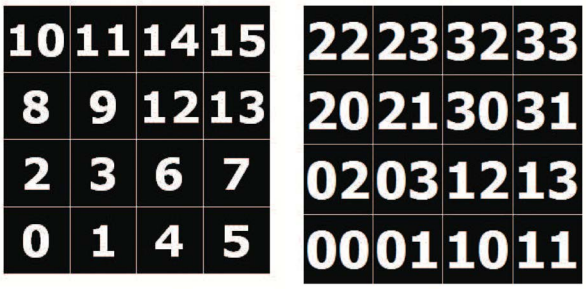


Figure 3. Hao's frame representation of $k = 2$.

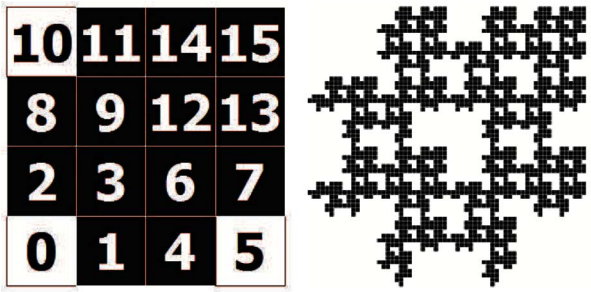


Figure 4. $B' = \{0,5,10\}$.

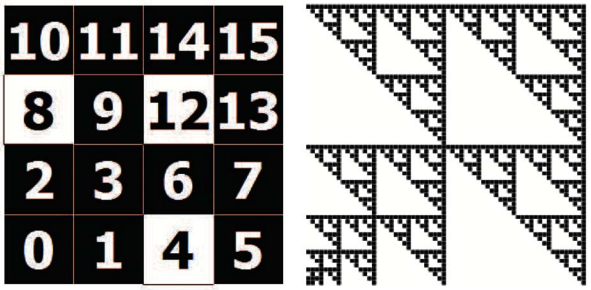


Figure 5. $B' = \{4,8,12\}$.

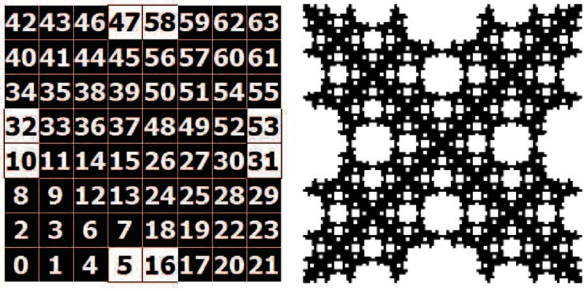


Figure 6. $B' = \{5,10,16,31,32,47,53,58\}$.

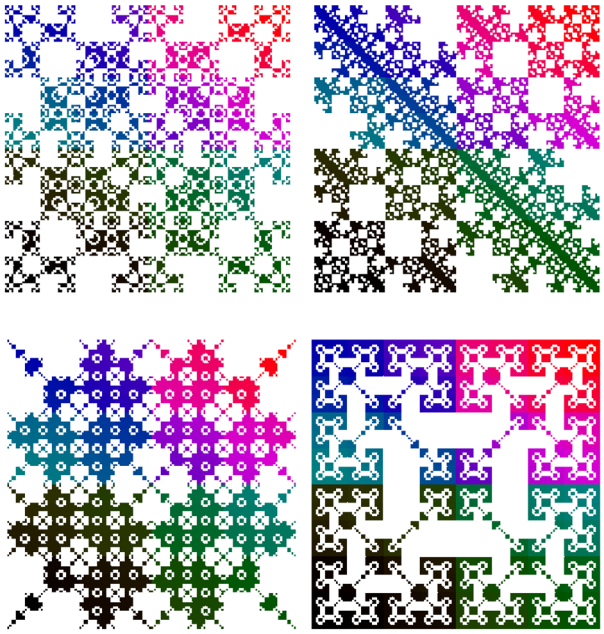


Figure 7. Other examples.

4. Conclusion

We established relations between the generalized Cantor sets and some DNA sequences with missing words. And we have associated Hao's frame representations and the generalized Sierpinski set with the generalized Cantor sets. The authors are interested in applying the analytical representation method to study the graphical results of space filling research works (cf. [23] [24] [25]).

Conflicts of Interest

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

References

[1] Mircea Anitas, E. and Slyamov, A. (2017) Structural Characterization of Chaos Game Fractals Using Small-Angle Scattering Analysis. *PLoS ONE*, **12**, e0181385. https://doi.org/10.1371/journal.pone.0181385

[2] Cattani, C. and Pierro, G. (2013) on the Fractal Geometry of DNA by the Binary Image Analysis. *Bulletin of Mathematical Biology*, **75**, 1544-1570. https://doi.org/10.1007/s11538-013-9859-9

[3] Badea, A.F., *et al.* (2013) Fractal Analysis of Elastographic Images for Automatic Detection of Diffuse Diseases of Salivary Glands: Preliminary Results. *Computational and Mathematical Methods in Medicine*, **2013**, Article ID: 347238. https://doi.org/10.1155/2013/347238

[4] Elloumi, M. and Zomaya, A.Y. (2014) Biological Knowledge Discovery Handbook. John Wiley & Sons, Hoboken, NJ. https://doi.org/10.1002/9781118617151

[5] Albrecht-Buehler, G. (2012) Fractal Genome Sequences. *Gene*, **498**, 20-27. https://doi.org/10.1016/j.gene.2012.01.090

[6] Ainsworth, C. (2009) Cells Go Fractal. *Nature-International Weekly Journal of Science*.

[7] Falconer, K.J. (1998) Techniques in Fractal Geometry. John Wiley & Sons, Hoboken, NJ.

[8] Hao, B.-L., Xie, H.-M. and Chen, G.-Y. (2000) Factorizable Language: From Dynamics to Bacterial Complete Genomes. *Physica A*: *Statistical Mechanics and Its Applications*, **288**, 10-20. https://doi.org/10.1016/S0378-4371(00)00411-8

[9] Hao, B.-L., Xie, H.-M., Chen, G.-Y. and Chen, G.-Y. (2000) Avoided Strings in Bacterial Complete Genomes and a Related Combinatorial Problem. *Annals of Combinatorics*, **4**, 247-255. https://doi.org/10.1007/PL00001279

[10] Yu, Z.-G., Hao, B.-L., Xie, H.-M. and Chen, G.-Y. (2000) Dimensions of Fractals Related to Languages Defined by Tagged Strings in Complete Genomes. *Chaos, Solitons & Fractals*, **11**, 2215-2222. https://doi.org/10.1016/S0960-0779(99)00141-1

[11] Crilly, A.J., Earnshaw, R. and Jones, H. (2013) Applications of Fractals and Chaos: The Shape of Things. Springer-Verlag, Berlin, Heidelberg.

[12] Huang, C.X. and Peng, S.L. (2006) A Note on Fractals of One Forbidden Word and Their Box Dimensions. *Fractals*, **14**, 327-337. https://doi.org/10.1142/S0218348X06003325

[13] Huang, C.X. and Peng, S.L. (2008) Fractals of Forbidden Words and Approximating Their Box Dimensions. *Physica A*: *Statistical Mechanics and Its Applications*, **387**, 703-716. https://doi.org/10.1016/j.physa.2007.09.009

[14] Jeffrey, H.J. (1990) Chaos Game Representation of Genetic Sequences. *Nucleic Acids Research*, **18**, 2163-2170. https://doi.org/10.1093/nar/18.8.2163

[15] Jeffrey, H.J. (1992) Chaos Game Visualization of Sequences. *Computers & Graphics*, **16**, 25-33. https://doi.org/10.1016/0097-8493(92)90067-6

[16] Tiňo, P. (1999) Spatial Representation of Symbolic Sequences Iterative Function Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*: *Systems and Humans*, **29**, 386-393. https://doi.org/10.1109/3468.769757

[17] Tiňo, P. (2002) Multifractal Properties of Hao's Geometric Representations of DNA Sequences. *Physica A*: *Statistical Mechanics and Its Applications*, **304**, 480-494. https://doi.org/10.1016/S0378-4371(01)00574-X

[18] Barnsley, M.F. (1988) Fractals Everywhere. Academic Press, Cambridge, MA.

[19] Berthelsen, C.L., Glazier, J.A. and Skolnik, M.H. (1992) Global Fractal Dimension of Human DNA Sequences Treated as Pseudorandom Walks. *Physical Review A*, **45**, 8902-8913. https://doi.org/10.1103/PhysRevA.45.8902

[20] Stanley, H.E., Buldyrev, S.V., Goldberger, A.L., Goldberger, Z.D., Havlin, S., Mantegna, R.N., Ossadnik, S.M., Peng, C.K. and Simons, M. (1994) Statistical Mechanics in Biology: How Ubiquitous Are Long-Range Correlations? *Physica A*: *Statistical Mechanics and Its Applications*, **205**, 214-253. https://doi.org/10.1016/0378-4371(94)90502-9

[21] Leong, P.M. and Morgenthaler, S. (1995) Random Walk and Gap Plots of DNA Sequences. *Bioinformatics*, **11**, 503-507. https://doi.org/10.1093/bioinformatics/11.5.503

[22] Solovyev, V.V., Korolev, S.V. and Lim, H.A. (1993) A New Approach for the Classification of Functional Regions of DNA Sequences Based of Fractal Representation. *International Journal of Genomic Research*, **1**, 109-128.

[23] Bagga, S., Girdhar, A., Trivedi, M.C. and Yang, Y.Z. (2016) RMI Approach to Cluster Based Cache Oblivious Peano Curves. 2016 *Second International Conference on*

Computational Intelligence & Communication Technology, Ghaziabad, India, 12-13 February 2016, 89-95. https://doi.org/10.1109/CICT.2016.26

[24] Makarov, B.M. and Podkorytov, A.N. (2017) On the Coordinate Functions of Peano Curves. *St. Petersburg Mathematical Journal*, **28**, 115-125. https://doi.org/10.1090/spmj/1441

[25] Platos, J., Kromer, P. and Snasel, V. (2015) Efficient Area Association Using Space Filling Curves. 2015 *International Conference on Intelligent Networking and Collaborative Systems*, Taipei, 2-4 September 2015, 322-326. https://doi.org/10.1109/INCoS.2015.39

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1697-1705
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Secular Effect of Geomagnetic Field and Gravitational Waves on Earth's Satellite Orbits

M. H. A. Youssef

Astronomy Department, Faculty of Science, Cairo University, Giza, Egypt
Email: mhyoussef@sci.cu.edu.eg

How to cite this paper: Youssef, M.H.A. (2019) Secular Effect of Geomagnetic Field and Gravitational Waves on Earth's Satellite Orbits. *Journal of Applied Mathematics and Physics*, **7**, 1697-1705.
https://doi.org/10.4236/jamp.2019.78116

Received: April 3, 2018
Accepted: August 10, 2019
Published: August 13, 2019

Abstract

In this work we study the perturbation and the change in the orbital elements due to the earth's magnetic field and the gravitational waves. The acceleration components are derived in the radial, transverse to it and normal to the orbital plane. The equation for the rates of variation of the elements is formed and solved to find the secular variation in the element for polar and equatorial satellites.

Keywords

Gravitational Waves, Earth's Magnetic Field, Perturbations, Orbital Mechanics

1. Introduction

One of the forces acting on an artificial satellite is the Earth's magnetic field (Roy, 1982) [1]. If the satellite has metal in its construction the Earth's magnetic field induces eddy currents in the satellite; in addition, a slight retardation acts on the satellite. The changes in the orbit due to this force are small; in any real situation the effect of the small force is to cause small departures from Keplerian motion, but these deviations can hardly give rise to large-scale changes of the effect of another force, like as in our work the force arising from gravitational waves. To gain concept of the concentration of charged particles in space that might lead to electromagnetic forces on a satellite, it is useful to estimate the flux and energy of those particles in the field of the Earth associated the aurorally zone, solar flares, solar winds, cosmic rays, and the inner and outer Van Allen radiation belts, and recently the gravitational waves. The question of the effect of such particles and the associated magnetic fields on satellite dynamics is yet to be resolved. As noted by (Bourdeau *et al.*, 1961) the motion of a satellite through a magnetic field B, produces an induced potential that is a function of position

on the satellite surface; that is, a function of a vector $\boldsymbol{R}$ from the satellite center to any point on the satellite surface [2]. The magnetic field strongly affects the motion of charged particles and the gravitational waves interact with magnetic field producing other effects on the motion of an orbiter in the Earth's gravitational field. The influence of the geo-magnetic field in satellite charging is quite important and should be taken into account in any detailed treatment [3]-[8]. In this work we investigate the perturbation on the elements of the satellite's orbit moving with velocity V through a magnetic field and existing of gravitational waves, considering the propagation of the gravitational waves in the same direction of the magnetic field.

2. The Acceleration Components of Magnetic field

A force arises from the interaction of the Earth's magnetic field and any electric charge of an orbiting body given by (Gelying and Westerman, 1971) [9]

$$\boldsymbol{F}_e = q_e \boldsymbol{V} \times \boldsymbol{B} \tag{1}$$

where q_e the charge is acquired by the satellite, $\boldsymbol{V}$ its velocity and $\boldsymbol{B}$ is the magnetic induction of the Earth's field. In MKS units

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \tag{2}$$

where the vector potential is

$$\boldsymbol{A} = \frac{\mu_0}{4\pi} \boldsymbol{M}_e \times \nabla\left(\frac{1}{r}\right) \tag{3}$$

$\boldsymbol{M}_e$ is the magnetic moment of the Earth, and μ_0 is the permeability of free space, carrying out the operations indicated in (1) through (3), we obtain

$$\frac{|\boldsymbol{F}_e|}{m} = q_e \frac{\mu_0 M_e}{4\pi m r^2} \frac{v}{r} \tag{4}$$

where m is the mass of satellite. As a first step to analyze effect of the magnetic field on the Kaplerian elements (a, e, i, ω, Ω, τ), it is necessary to represent the acceleration exerted by the magnetic field on a satellite as a function of the radial, orthogonal, and normal, perturbing components (Baker, 1967) [10]. This can be done in terms of either the true anomaly f, or eccentric anomaly E. The resolving components of the disturbing force in the direction of the satellite radius-vector S, tangential to the orbit T and perpendicular to it W, are given as

$$\begin{aligned} S &= q_e \frac{\mu_0 M_e}{4\pi m} \sqrt{\frac{\mu}{P^7}} e \sin f \left(1 + e\cos f\right)^3 \\ T &= q_e \frac{\mu_0 M_e}{4\pi m} \sqrt{\frac{\mu}{P^7}} \left(1 + e\cos f\right)^4 \\ W &= 0 \end{aligned} \tag{5}$$

where $P = a\left(1 - e^2\right)$ is the parameter, a is the semi-major axis, e is the eccentricity of the orbit, μ is the Earth's gravitational constant and f is the true anomaly.

3. The Change of the Orbital Elements

According to Gauss form of Lagrange planetary equations under the action of the geomagnetic field and normal incident of gravitational waves the changes of the osculating elements for an elliptical orbit have the form

$$\begin{aligned}
\frac{\mathrm{d}a}{\mathrm{d}t} &= \frac{2}{n\sqrt{1-e^2}}\left\{e\left(S_{mg}+S_{gw}\right)\cos f+\frac{p}{r}\left(T_{mg}+T_{gw}\right)\right\}\\
\frac{\mathrm{d}e}{\mathrm{d}t} &= \frac{\sqrt{1-e^2}}{na}\left\{\sin f\left(S_{mg}+S_{gw}\right)+\left(\cos f+\cos E\right)\left(T_{mg}+T_{gw}\right)\right\}\\
\frac{\mathrm{d}i}{\mathrm{d}t} &= \frac{r\cos\left(f+\omega\right)}{na^2\sqrt{1-e^2}\sin i}\left(W_{mg}+W_{gw}\right)\\
\frac{\mathrm{d}\Omega}{\mathrm{d}t} &= \frac{r\sin\left(f+\omega\right)}{na^2\sqrt{1-e^2}\sin i}\left(W_{mg}+W_{gw}\right)\\
\frac{\mathrm{d}\omega}{\mathrm{d}t} &= \frac{\sqrt{1-e^2}}{nae}\left\{-\cos f\left(S_{mg}+S_{gw}\right)+\left(1+\frac{r}{p}\right)\sin f\left(T_{mg}+T_{gw}\right)\right\}\\
&\quad -\frac{r\sin\left(f+\omega\right)}{na^2\sqrt{1-e^2}}\cot i\left(W_{mg}+W_{gw}\right)
\end{aligned} \tag{6}$$

where (S_{mg}, T_{mg}, W_{mg}) and (S_{gw}, T_{gw}, W_{gw}) are referred to the acceleration components for the magnetic field and gravitational waves in the radial direction S, the transverse direction T at right angle to S in the orbital plane and perpendicular to the orbital plane W as in Figure 1.

Where Ω, ω, $u=\left(f+\omega\right)$ and i are the angles of longitude of ascending node, argument of perigee, true anomaly and inclination respectively. (S_{mg}, T_{mg}, W_{mg}) are the resolving acceleration components of geomagnetic force which derived in Equation (5). The resolving acceleration components of gravitational waves force are

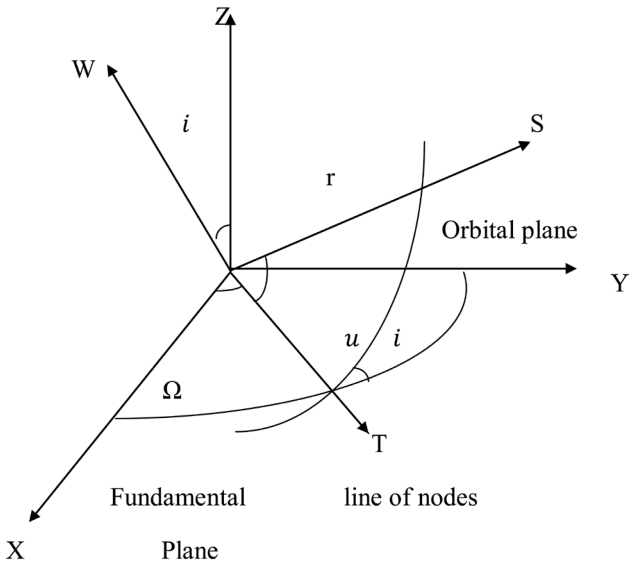


Figure 1. The disturbing force of gravitational waves in (S, T, W) directions.

$$S_{gw} = P_x F_x + P_y F_y$$
$$T_{gw} = Q_x F_x + Q_y F_y$$
$$W_{gw} = W_x F_x + W_y F_y \tag{7}$$

where

$$F_x = h_1 x + h_2 y$$
$$F_y = h_2 x - h_1 y$$
$$F_z = 0 \tag{8}$$

F_x, F_y and F_z are the components of the acceleration of normal ancient gravitational waves in (x, y, z) coordinates and

$$h_1 = \frac{1}{2}\frac{\partial^2 h_{11}}{\partial t^2};\quad h_2 = \frac{1}{2}\frac{\partial^2 h_{12}}{\partial t^2} \tag{9}$$
$$h_{11} = h_+ \cos\left(n_g t + \alpha_1\right)$$
$$h_{12} = h_\times \cos\left(n_g t + \alpha_2\right) \tag{10}$$

where n_g is the frequency of the wave, α_1 and α_2 are the phase difference, h_+ and $h_\times$ are the amplitude of the wave in the two orthogonal directions in the transverse plane [11]. Therefore

$$h_1 = -\frac{1}{2} n_g^2 h_+ \cos\left(n_g t + \alpha_1\right)$$
$$h_2 = -\frac{1}{2} n_g^2 h_\times \cos\left(n_g t + \alpha_2\right) \tag{11}$$

We have $P^\wedge, Q^\wedge$ and $W^\wedge$ are the unit vectors in the direction of $\boldsymbol{r}$, normal to $\boldsymbol{r}$ in the orbital plane and normal to the orbital plane respectively,

$$P_x = \cos\Omega\cos u - \sin\Omega\sin u\cos i$$
$$P_y = \sin\Omega\cos u + \cos\Omega\sin u\cos i$$
$$P_z = \sin u\sin i$$
$$Q_x = -\cos\Omega\sin u - \sin\Omega\cos u\cos i$$
$$Q_y = -\sin\Omega\sin u + \cos\Omega\cos u\cos i$$
$$Q_z = \cos u\sin i$$
$$W_x = \sin\Omega\sin i$$
$$W_y = -\cos\Omega\sin i$$
$$W_z = \cos i \tag{12}$$

It follows that

$$x = r\left(\cos\Omega\cos u - \sin\Omega\sin u\cos i\right)$$
$$y = r\left(\sin\Omega\cos u + \cos\Omega\sin u\cos i\right)$$
$$z = \sin u\sin i \tag{13}$$

Therefore from Equations ((8) to (13)) Equation (7) yield to

$$
\begin{aligned}
S_{gw} &= r\left\{A_1 + A_2\cos 2(u) + A_3\sin 2(u)\right\}\\
T_{gw} &= r\left\{-A_2\sin 2(u) + A_3\cos 2(u)\right\}\\
W_{gw} &= r\left\{C_1\cos(u) + C_2\sin(u)\right\}\sin i
\end{aligned}
\tag{14}
$$

where

$$
\begin{aligned}
A_1 &= \frac{1}{2}\sin^2 i\left\{h_1\cos 2\Omega + h_2\sin 2\Omega\right\}\\
A_2 &= \frac{1+\cos^2 i}{2}\left\{h_1\cos 2\Omega + h_2\sin 2\Omega\right\}\\
A_3 &= \cos i\left\{-h_1\sin 2\Omega + h_2\cos 2\Omega\right\}\\
C_1 &= h_1\sin 2\Omega - h_2\cos 2\Omega\\
C_2 &= \cos i\left\{h_1\cos 2\Omega + h_2\sin 2\Omega\right\}
\end{aligned}
\tag{15}
$$

For the approximate integration of Equations (6), we expand all the functions of the orbital coordinates on the right sides in series in powers of e. The coefficients in these series will be trigonometric functions of the mean anomaly

$$
\begin{aligned}
M &= n(t-\tau)\\
\cos kf &= \cos kM + ke\left[\cos(k+1)M - \cos(k-1)M\right]\\
\sin kf &= \sin kM + ke\left[\sin(k+1)M - \sin(k-1)M\right];\quad K = 1,2,3,\cdots\\
\cos E &= \cos M + \frac{e}{2}(\cos 2M - 1)\\
r &= a(1 - e\cos M)
\end{aligned}
\tag{16}
$$

where n the mean angular velocity of satellite, t the initial time of the motion, and τ is the time of perigee passage. Without loss of generality, we can take the satellite lies on the line of nodes at the initial time, so $\omega = \tau = 0$ and Equations (6) become in the form

$$
\begin{aligned}
\frac{\mathrm{d}a}{\mathrm{d}t} &= \frac{2a}{n}\left\{T_0 + e\left(T_1 + S_0\sin M + T_0\cos M\right)\right\}\\
\frac{\mathrm{d}e}{\mathrm{d}t} &= \frac{1}{n}\left\{S_0\sin M + 2T_0\cos M + e\left[S_1\sin M + 2T_1\cos M\right.\right.\\
&\qquad \left.\left. + S_0\sin 2M + \frac{3}{2}T_0(\cos 2M - 1)\right]\right\}\\
\frac{\mathrm{d}i}{\mathrm{d}t} &= \frac{\sin i}{n}\left\{W_0\cos M + e\left[\frac{W_0}{2}(\cos 2M - 3) + W_1\cos M\right]\right\}\\
\frac{\mathrm{d}\Omega}{\mathrm{d}t} &= \frac{1}{n}\left\{W_0\sin M + e\left[W_1\sin M + \frac{1}{2}W_0\sin 2M\right]\right\}\\
\frac{\mathrm{d}\omega}{\mathrm{d}t} &= \frac{1}{ne}\left\{-S_0\cos M + 2T_0\sin M + e\left[S_0 - S_1\cos M + 2T_1\sin M - S_0\cos 2M\right.\right.\\
&\qquad \left.\left. + \frac{3}{2}T_0\sin 2M\right]\right\} - \frac{\cos i}{n}\left\{W_0\sin M + e\left[W_1\sin M + \frac{1}{2}W_0\sin 2M\right]\right\}
\end{aligned}
\tag{17}
$$

where

$$S_0 = A_1 + A_2\cos 2M + A_3\sin 2M + q_e\frac{\mu_0 M_e}{4\pi m}\sqrt{\frac{\mu}{P^7}}e\sin M$$

$$S_1 = -A_1\cos M + \frac{1}{2}A_2\left(3\cos 3M - 5\cos M\right) + \frac{1}{2}A_3\left(3\sin 3M - 5\sin M\right)$$

$$T_0 = -A_2\sin 2M + A_3\cos 2M + q_e\frac{\mu_0 M_e}{4\pi m}\sqrt{\frac{\mu}{P^7}}\left(1 + 4e\cos M\right)$$

$$T_1 = -\frac{1}{2}A_2\left(3\sin 3M - 5\sin M\right) + \frac{1}{2}A_3\left(3\cos 3M - 5\cos M\right)$$

$$W_0 = C_1\cos M + C_2\sin M$$

$$W_1 = \frac{1}{2}\left\{C_1\left(\cos 2M - 3\right) + C_2\sin 2M\right\} \tag{18}$$

Because of the smallness of the eccentricity *e*, it is of interest to consider such effects only for small powers of *e*. Therefore Equation (17) is represented the first order approximations of Equation (6) and the effect of geomagnetic and GW on the orbital elements. A general analysis of Equation (17) shows that the semi-major axis *a* and the position of the orbital plane in space determined by the angles Ω and *i*, change for a wave frequency, wave amplitude and mass ratio of satellite $\frac{q}{m}$ due to the induction of Earth's magnetic field. For $i = \frac{\pi}{2}$, the plane of orbit of which is parallel to the direction of the gravitational waves and the direction of magnetic field given by the quantities $h_\times = 0$ and $\alpha_1 = \pm\frac{\pi}{2}$, in this case only the longitude of ascending node Ω changes and the shape of the orbit is constant. If the initial data for the orbit and the waves are defined by $i = \frac{\pi}{2}$, $\Omega = 0$, $h_\times = 0$ and $\alpha_1 = \pm\frac{\pi}{2}$ then the position of the orbital plane is constant in space and only the semi-major axis *a* changes without a change in the other parameters. For the initial data $i = \frac{\pi}{2}$, $h_\times = \alpha_1 = 0$ and $\Omega = \frac{\pi}{4}$ there is only a deviation of the orbital plane from the wave direction, determined by the angle *i*. For the case $i = h_\times = 0$ and $\alpha_1 = \pm\frac{\pi}{2}$ only eccentricity changes. In case $i = h_\times = 0$ and $\alpha_1 = 0$, this time only the angular distance ω changes.

4. Numerical Simulation

A numerical simulation is developed to test several of the above results. The simulation is a Runge-Kutta 4th order integration of system defined by Equations (17) with respect to the mean anomaly *M* performed by MATHEMATICA V10. The simulation is valid for any orbit for charged satellite in a non-tilted dipole field with GW. The results of two different situations are presented here for polar and equatorial orbits. Considering the period of satellite is four hours and with semi-major 12,600 km. The effects of the magnetic field and the gravitational

waves are proportional to the charge to mass ratio $\frac{q}{m}$ and the frequency of the gravitational waves. Considering the wave's frequency n_g of the same order of magnitude of the magnetic field 10^{-9}, wave's amplitude of order 10^{-21} as coming from the bursts sources, and the mass ratio $\frac{q}{m} = 0.0393\,\mathrm{C/kg}$ for satellite has mass 300 kg and $q = 117\,\mathrm{C}$. This means that with two parameters we determined the perturbations on the orbital motion and we can control on it. Also due to the interaction of GW with magnetic field the variation on the orbital element will periodic with the time.

Figure 2 displays the variation in the orbital plane due the perturbation of Ω in radian for polar orbits with inclination $i = 90^\circ$, $\omega = 0^\circ$ or 180° or 360° and $\Omega = \frac{\pi}{4}$.

Figure 3 displays the variation of the semi-major a in radians for the polar circular orbit with $\Omega = 0^\circ$, $h_\times = 0$ and $\alpha_1 = \pm\frac{\pi}{2}$.

Figure 4 displays the variation of the inclination i in radian for the polar circular orbit with $\Omega = \frac{\pi}{4}$, $h_\times = 0$ and $\alpha_1 = 0^\circ$.

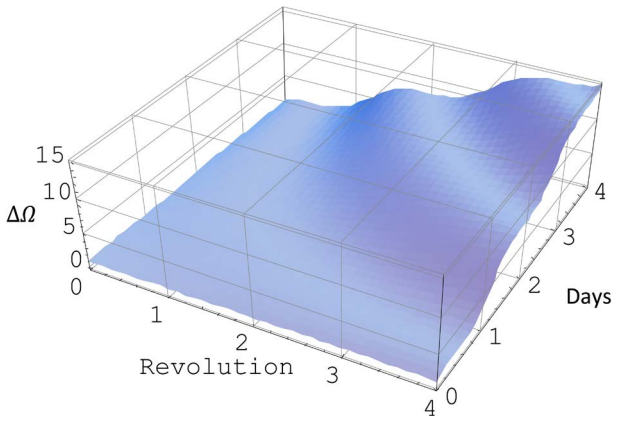


Figure 2. The variations of the orbital plane by angle Ω for polar orbit.

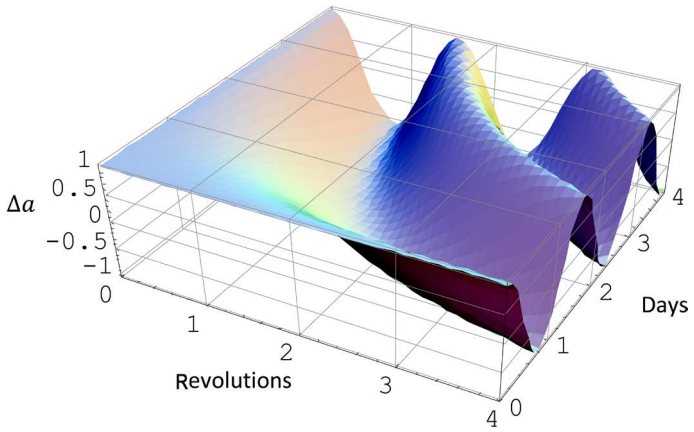


Figure 3. The variations of the semi-major a for polar orbit.

Figure 5 displays the variation of the eccentricity e for equatorial orbit with the case $i = h_x = 0$ and $\alpha_1 = \pm\frac{\pi}{2}$.

Figure 6 displays the variation of the angular distance ω in radian for equatorial orbit with the case $i = h_x = 0$ and $\alpha_1 = 0^\circ$.

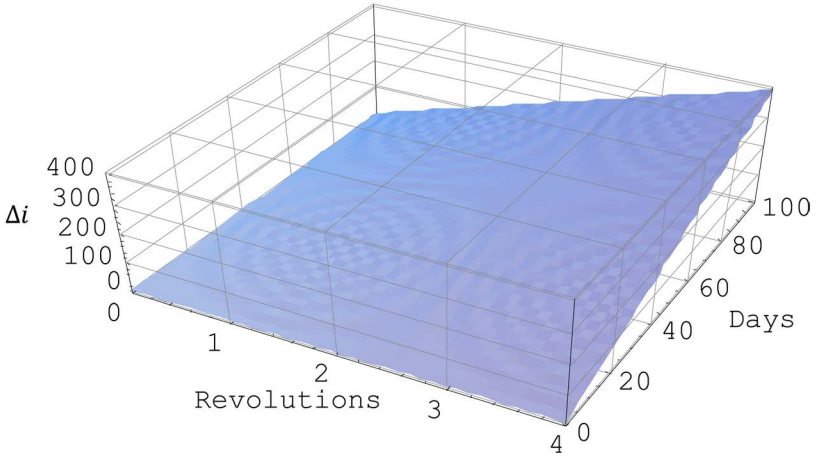


Figure 4. The variations of the inclination i for polar orbit.

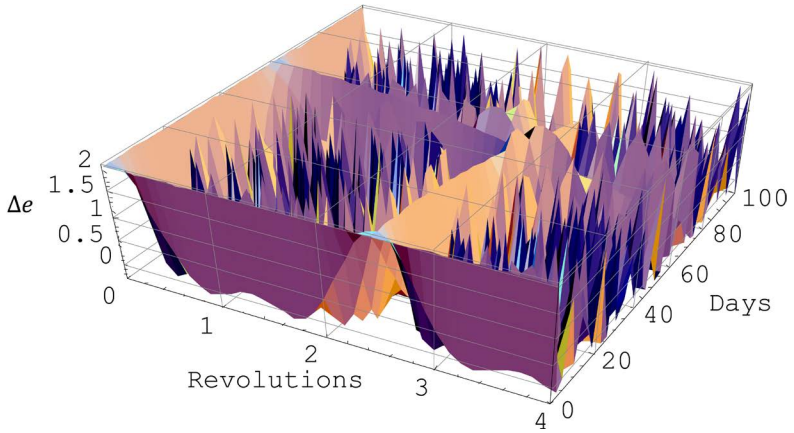


Figure 5. The variations of the eccentricity e for equatorial orbit.

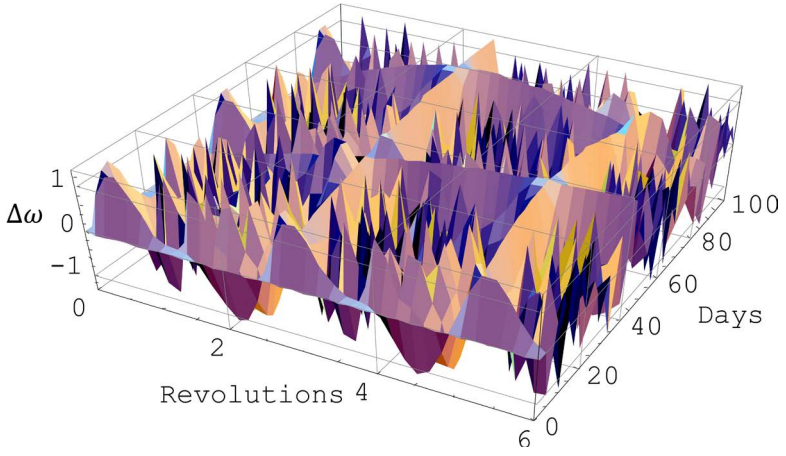


Figure 6. The variations of the angular distance ω for equatorial orbit.

5. Discussion and Conclusion

Summarizing the results we conclude that there is secular effect due to geomagnetic and gravitational waves on the orbital plane of polar orbits (the longitude of node Ω and inclination i) and on the shape (the semi-major axis a) according to phase and phase difference of gravitational waves. For equatorial orbits the secular effect will on the size on the orbit (the eccentricity e and the argument of perigee ω). The amount of the variation of the orbital elements depends on the direction of propagation of gravitational waves and on the frequency and its phase difference which changes the amount of charges on the satellite and consequently the perturbations on the orbital elements. This effect is small but it is important for studying the effect of GW with the magnetic field of the Earth during interval of time and their effects on artificial satellites.

Conflicts of Interest

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

References

[1] Roy, A.E. (1982) Orbital Motion. Adam Higher, Bristol.

[2] Bourdeau, R., Donley, J., Serbu, G. and Whipple, E. (1961) Measurements of Sheath Currents and Equilibrium Potential on the Explorer VIII Satellite. AAS, San Diego, 61-63.

[3] Bourdeau, R. (1963) On the Interaction between a Spacecraft and an Ionized Medium. *Space Science Reviews*, **1**, 719-728. https://doi.org/10.1007/BF00212449

[4] Jefimenco, O. (1959) Effect of the Earth's Magnetic Field on the Motion of an Artificial Satellite. *American Journal of Physics*, **27**, 344-348. https://doi.org/10.1119/1.1934848

[5] Brudin, C. (1963) Effects of Charged Particles on the Motion of an Artificial Satellite. *AIAA Journal*, **1**, 2529-2538. https://doi.org/10.2514/3.2105

[6] Vokrouhlicky, D. (1990) Lorentz Force Perturbations of the Orbit of an Electrically Charged Satellite—Case of Varying Charge. *Bulletin of Astronomy and Insturment Czech.*, **41**, 205-211.

[7] Vokrouhlicky, D. (1989) The Geomagnetic Effects on the Motion of an Electrically Charged Artificial Satellite. *Celestial Mechanics and Dynamical Astronomy*, **46**, 85-104. https://doi.org/10.1007/BF02426715

[8] Streetman, B. and Peck, M.A. (2006) New Synchronous Orbits Using the Geomagnetic Lorentz Force. *AIAA Guidance, Navigation, and Control Conference*, Keystone, 21-24 August 2006, 1677. https://doi.org/10.2514/6.2006-6804

[9] Geyling, F. and Westerman, H. (1971) Introduction to Orbital Mechanics. Addison-Wesley, Boston.

[10] Baker, R. (1967) Astrodynamics, Applications and Advanced Topics. Academic Press, New York.

[11] Straumann, N. (1984) General Relativity and Relativistic Astrophysics. Springer-Verlag, Berlin. https://doi.org/10.1007/978-3-642-84439-3

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1706-1717
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Global Analysis of an SEIR Epidemic Model with Infectious Force under Intervention Strategies

Minmin Zhou, Tiansi Zhang

College of Science, University of Shanghai for Science and Technology, Shanghai, China
Email: 1447512612@qq.com

How to cite this paper: Zhou, M.M. and Zhang, T.S. (2019) Global Analysis of an SEIR Epidemic Model with Infectious Force under Intervention Strategies. *Journal of Applied Mathematics and Physics*, **7**, 1706-1717.
https://doi.org/10.4236/jamp.2019.78117

Received: July 21, 2019
Accepted: August 13, 2019
Published: August 16, 2019

Abstract

In this paper, we investigate the global stability of an SEIR (Susceptible-Exposed-Infected-Remove) epidemic model with infectious force under intervention strategies. To address this issue, we prove that the basic reproduction number R_0 plays an essential role in determining whether the disease extincts or persists. If $R_0 \le 1$, there is a unique disease-free equilibrium point of the model which is globally asymptotically stable and the disease dies out, and if $R_0 > 1$, there exists a unique endemic equilibrium point which is globally asymptotically stable and the disease persists.

Keywords

SEIR Epidemic Model, Intervention Strategies, Basic Reproduction Number, Global Stability

1. Introduction

As we all know, infectious disease has been ravaging human beings for thousands of years. In the long history, infectious disease has brought many disasters to human beings. For a long time, people have been fighting various infectious disease, and many methods have been used to study the spread of infectious disease, such as to control and eliminate infectious disease. Since the pioneer work of Kermack and McKendrick [1], mathematical models have been contributing to improve our understanding of infectious disease dynamics and helping us develop preventive measures to control infection spread qualitatively and quantitatively [2] [3]. The effect of intervention strategies, such as border screening, mask wearing, quarantine, isolation or communications through the mass media,

plays an important role in administering efficient interventions to control disease spread and hopefully eliminate epidemic disease [4]-[9].

In recent years, a number of mathematical models have been formulated to describe the impact of intervention strategies on the dynamics of infectious disease [4]-[11]. Tang and Xiao indicate that strict interventions have been taken in mainland of China to slow down the initial spread of the disease [10] [12], and awareness through media and education plays an important role in changing behavior or contact patterns, and hence in limiting the spread of infectious disease [9].

In particular, Wang [4] formulates and analyzes an SIRS (Susceptible-Infected-Remove-Susceptible) epidemic model to study the impact of intervention strategies on the spread infectious disease and find that intervention strategies decrease endemic levels and tend to simplify diseases. In this article, we analyze the SEIR epidemic model to study the impact of intervention strategies on the spread of infectious disease.

In real life, epidemic tends to have an incubation period, as susceptible to infection after contacting with infected people. First of all, carrying virus, the virus is not immediately, but after a period of time, to onset and into the herd of infected people. In this paper, we mainly focus on the global stability analysis of the steady states for an SEIR epidemic model with infectious force under intervention.

We consider the global properties of this SEIR model and show that if the basic reproduction number $R_0 \le 1$, the disease-free equilibrium point is globally asymptotically stable, while if $R_0 > 1$, the disease-free equilibrium point is unstable and the unique endemic equilibrium point is globally asymptotically stable.

The rest organization of this article is arranged as follows: In Section 2, we present the model. In Section 3, we illustrate the main results and proof the main results in details. In Section 4, we provide the application of the results to SEIR model with infection force under intervention policy to support our findings. In the last section, we provide a brief discussion and summary of the results.

2. Model Derivations

We propose a deterministic SEIR epidemic model with infectious force. The model is given by

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = A - \mu S - H(I)S, \\ \dfrac{\mathrm{d}E}{\mathrm{d}t} = H(I)S - (\mu + q)E, \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = qE - (\mu + \delta + \gamma)I, \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = \gamma I - \mu R, \end{cases} \tag{1}$$

where S, E, I and R denote the number of susceptible, exposed, infective and recovered individuals at time t, respectively, and $N(t)=S(t)+E(t)+I(t)+R(t)$. All parameters are positive with

A: the recruitment rate of the population;

μ: the nature death rate of the population;

q: the constant rate such that the exposed individuals become infective;

δ: the disease inducing death rate;

γ: the natural recovery rate of the infective individuals.

The infective force $H(I)$ in (1) is a function of infective individuals which plays a key role in determining the transmission of disease. There are several different nonlinear transmission functions proposed by researchers, see more details, we refer to [4] [5] [13] [14] and the reference therein.

Model (1) includes the adaption of individuals behavior under intervention polices. For example, $H(I)$ may decrease as the number of infective individuals increase due to the fact that the population may tend to reduce the number of contacts per unit time under intervention polices. This has been interpreted as the psychological effect [5]. Mathematically, this phenomenon can be modeled as the infection force $H(I)$ which is increasing when I is small and decreasing when I is large. For simplicity in notations, we suppose that the infection force $H(I)$ can be factorized into $\frac{\beta I}{f(I)}$, where $\frac{1}{f(I)}$ represents the effect of intervention strategies on the reduction of valid contact coefficient β [4]. It is worthy to note that, in the absence of intervention strategies, *i.e.* $f(I)=1$, the incidence rate becomes the well-known bilinear transmission rate βSI. To ensure a nonmonotonic infection force, we make the following assumptions [2]:

(H1) $f(0)>0$ and $f'(I)>0$ for $I>0$.

(H2) There is ζ such that $\left(\frac{I}{f(I)}\right)'>0$ for $0<I<\zeta$ and $\left(\frac{I}{f(I)}\right)'<0$ for $I>\zeta$.

In epidemiology, these assumptions describe the effects of intervention strategies determining by a critical value ζ: if $0<I<\zeta$, the incidence rate is increasing, while if $I>\zeta$, the incidence rate is decreasing. Thus we can establish the following SEIR epidemic model:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t}=A-\mu S-\dfrac{\beta I}{f(I)}S, \\ \dfrac{\mathrm{d}E}{\mathrm{d}t}=\dfrac{\beta I}{f(I)}S-(\mu+q)E, \\ \dfrac{\mathrm{d}I}{\mathrm{d}t}=qE-(\mu+\delta+\gamma)I, \\ \dfrac{\mathrm{d}R}{\mathrm{d}t}=\gamma I-\mu R. \end{cases} \tag{2}$$

Since R does not appear in the first three equations of system (2), it can be

reduced to the following three-dimensional system

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = A - \mu S - \dfrac{\beta I}{f(I)} S, \\ \dfrac{\mathrm{d}E}{\mathrm{d}t} = \dfrac{\beta I}{f(I)} S - (\mu + q) E, \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = qE - (\mu + \delta + \gamma) I. \end{cases} \tag{3}$$

where the state space is in the first quadrant $R_+^3 = \{(S, E, I) : S > 0, E > 0, I > 0\}$.

It follows from system (3) that:

$$A - (\mu + \delta + \gamma) M \le \frac{\mathrm{d}M}{\mathrm{d}t} \le A - \mu M.$$

where $M = S + E + I$.

Hence, by integrating the above inequality, there is

$$\frac{A}{\mu + \delta + \gamma} + \left(M(0) - \frac{A}{\mu + \delta + \gamma} \right) \mathrm{e}^{-(\mu+\delta+\gamma)t} \le M \le \frac{A}{\mu} + \left(M(0) - \frac{A}{\mu} \right) \mathrm{e}^{-\mu t}.$$

Then let $t \to \infty$, we can get

$$\frac{A}{\mu + \delta + \gamma} \le \lim_{t\to\infty} \inf(M) \le \lim_{t\to\infty} \sup(M) \le \frac{A}{\mu}.$$

That is

$$\frac{A}{\mu + \delta + \gamma} \le S + E + I \le \frac{A}{\mu}.$$

The feasible region for system (3) is thus a bounded set Γ:

$$\Gamma = (S, E, I) \in R_+^3 : 0 < S + E + I \le \frac{A}{\mu} \subset R_+^3.$$

The region Γ is a positive invariant set for model (3). Moreover, every trajectory of model (3) eventually stays in a compact subset of Γ.

3. Main Results and Proof of Main Results

In this section, we give main results and the proof of main results.

Notice that model (3) has a disease-free equilibrium point $P_0 = \left(\dfrac{A}{\mu}, 0, 0 \right)$ for all parameter. A main concern of deterministic epidemic model is to find conditions when a disease introduced into a community can develop into a large outbreak, and if it does the disease may become endemic. A useful threshold in this regard for deterministic models is called basic reproduction number R_0.

Let $X = (E, I, S)^{\mathrm{T}}$, system (3) can be written as [15]

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X) - V(X),$$

where

$$F(X)=\begin{pmatrix}\dfrac{\beta IS}{f(I)}\\0\\0\end{pmatrix},V(X)=\begin{pmatrix}(\mu+q)E\\-qE+(\mu+\delta+\gamma)I\\-A+\mu S+\dfrac{\beta IS}{f(I)}\end{pmatrix}.$$

The jacobian matrices of (9) at the disease-free equilibrium point are

$$DF(P_0)=\begin{bmatrix}F&0\\0&0\end{bmatrix},DV(P_0)=\begin{bmatrix}V&0\\J_1&J_2\end{bmatrix},$$

where

$$F=\begin{bmatrix}0&\dfrac{\beta A}{\mu f(0)}\\0&0\end{bmatrix},V=\begin{bmatrix}\mu+q&0\\-q&\mu+\delta+\gamma\end{bmatrix}.$$

So the next generation matrix of model (3) is

$$FV^{-1}=\begin{bmatrix}\dfrac{\beta Aq}{\mu f(0)(\mu+q)(\mu+\delta+\gamma)}&\dfrac{\beta A}{\mu f(0)(\mu+\delta+\gamma)}\\0&0\end{bmatrix}.$$

and the spectral radius of FV^{-1} is

$$\rho\left(FV^{-1}\right)=\frac{\beta Aq}{\mu f(0)(\mu+q)(\mu+\delta+\gamma)}.$$

Therefore the basic reproduction number R_0 [16] is

$$R_0=\frac{\beta Aq}{\mu f(0)(\mu+q)(\mu+\delta+\gamma)}.$$

The epidemic model (3) has two equilibrium points: one is the disease-free equilibrium point $P_0=\left(\dfrac{A}{\mu},0,0\right)$ which exists for all parameter values; and the second is the endemic equilibrium point $P^*=\left(S^*,E^*,I^*\right)$ which is a positive solution of the following system

$$\begin{cases}A-\mu S^*-\dfrac{\beta I^*}{f\left(I^*\right)}S^*=0,\\[2ex]\dfrac{\beta I^*}{f\left(I^*\right)}S^*-(\mu+q)E^*=0,\\[2ex]qE^*-\left(\mu+\delta+\gamma\right)I^*=0.\end{cases}\tag{4}$$

From (4),

$$S^*=\frac{(\mu+q)(\mu+\delta+\gamma)}{q\beta}f\left(I^*\right),E^*=\frac{\mu+\delta+\gamma}{q}I^*.$$

and

$$A-\frac{(\mu+q)(\mu+\delta+\gamma)}{q}I^*-\frac{\mu(\mu+q)(\mu+\delta+\gamma)}{\beta q}f\left(I^*\right)=0.$$

Set

$$Q(I)=A-\frac{(\mu+q)(\mu+\delta+\gamma)}{q}I-\frac{\mu(\mu+q)(\mu+\delta+\gamma)}{\beta q}f(I).$$

It follows from the assumption (H1) that

$$Q'(I)=A-\frac{(\mu+q)(\mu+\delta+\gamma)}{q}-\frac{\mu(\mu+q)(\mu+\delta+\gamma)}{\beta q}f'(I)<0.$$

Hence $Q(I)$ is a decreasing function.

Note that

$$Q(0)=\frac{\mu(\mu+q)(\mu+\delta+\gamma)f(0)}{\beta q}(R_0-1).$$

If $R_0>1$, we can know that $Q(0)>0$, and because $Q(I)$ is a decreasing function so $Q(I)=0$ has a unique positive solution I^*, then model (3) has a unique endemic equilibrium $P^*=(S^*,E^*,I^*)$ with

$$S^*=\frac{(\mu+q)(\mu+\delta+\gamma)}{\beta q}f(I^*),E^*=\frac{\mu+\delta+\gamma}{q}I^*.$$

So we know that the unique endemic equilibrium point $P^*=(S^*,E^*,I^*)$ of model (3) exists when $R_0>1$.

Theorem 3.1. *The disease-free equilibrium point* $P_0=\left(\frac{A}{\mu},0,0\right)$ *of model* (3) *is globally asymptotically stable if* $R_0\le 1$ *or unstable if* $R_0>1$.

Proof. Define the Lyapunov function

$$V(S,E,I)=\frac{1}{2}\left(S-\frac{A}{\mu}\right)^2+\frac{A}{\mu}E+\frac{A(\mu+q)}{\mu q}I.$$

Take derivative of V along the solution of model (3), there is

$$\begin{aligned}\frac{\mathrm{d}V}{\mathrm{d}t}&=\left(S-\frac{A}{\mu}\right)\frac{\mathrm{d}S}{\mathrm{d}t}+\frac{A}{\mu}\frac{\mathrm{d}E}{\mathrm{d}t}+\frac{A(\mu+q)}{\mu q}\frac{\mathrm{d}I}{\mathrm{d}t}\\&=\left(S-\frac{A}{\mu}\right)\left(A-\mu S-\frac{\beta IS}{f(I)}\right)+\frac{A}{\mu}\left(\frac{\beta IS}{f(I)}-(\mu+q)E\right)\\&\quad+\frac{A(\mu+q)}{\mu q}\left(qE-(\mu+\delta+\gamma)I\right)\\&=-\mu\left(S-\frac{A}{\mu}\right)^2-\left(S-\frac{A}{\mu}\right)\frac{\beta IS}{f(I)}+\frac{A}{\mu}\frac{\beta IS}{f(I)}-\frac{A(\mu+q)(\mu+\delta+\gamma)}{\mu q}I\\&=-\left(\mu+\frac{\beta I}{f(I)}\right)\left(S-\frac{A}{\mu}\right)^2+\frac{A}{\mu}\frac{A\beta-\mu(\mu+q)(\mu+\delta+\gamma)\frac{f(I)}{q}}{f(I)}I.\end{aligned}$$

Note that $f(I)=f(0)+f'(0)I+o(I)$, we have

$$\begin{aligned}&A\beta-\mu(\mu+q)(\mu+\delta+\gamma)\frac{f(I)}{q}\\&=A\beta-\frac{\mu(\mu+q)(\mu+\delta+\gamma)}{q}(f(0)+f'(0)I+o(I))\end{aligned}$$

$$\le A\beta - \frac{\mu(\mu+q)(\mu+\delta+\gamma)}{q} f(0) - \frac{\mu(\mu+q)(\mu+\delta+\gamma)}{q} f'(0) I$$

$$\le -\frac{\mu(\mu+q)(\mu+\delta+\gamma) f(0)}{q}(1-R_0) - \frac{\mu(\mu+q)(\mu+\delta+\gamma)}{q} f'(0) I$$

$$\le 0.$$

when $R_0 \le 1$.

So

$$\frac{\mathrm{d}V}{\mathrm{d}t} \le -\left(\mu + \frac{\beta I}{f(I)}\right)\left(S - \frac{A}{\mu}\right)^2 - \frac{A}{\mu}\frac{\mu(\mu+q)(\mu+\delta+\gamma)}{q}\frac{f(0)(1-R_0)I + f'(0)I^2}{\mu f(I)} < 0.$$

By applying the Lyapunov-LaSall asymptotic stability theorem [17] [18], we conclude that P_0 is globally asymptotically stable if $R_0 \le 1$.

When $R_0 > 1$, the Jacobian matrix of model (3) evaluated at P_0 is

$$J(P_0) = \begin{pmatrix} -\mu & 0 & -\dfrac{\beta A}{\mu f(0)} \\ 0 & -\mu - q & \dfrac{\beta A}{\mu f(0)} \\ 0 & q & -(\mu+\delta+\gamma) \end{pmatrix},$$

which has an eigenvalue $-\mu < 0$. Denoted by

$$B = \begin{pmatrix} -\mu - q & \dfrac{\beta A}{\mu f(0)} \\ q & -\mu - \delta - \gamma \end{pmatrix}.$$

We find that when $R_0 > 1$, $\det B = (1-R_0)(\mu+q)(\mu+\delta+\gamma) < 0$, $trB = -(\mu+q+\mu+\delta+\gamma) < 0$, so the matrix B must have a positive eigenvalue. Thus, the disease-free equilibrium point P_0 is unstable whenever $R_0 > 1$. This ends the proof.

Theorem 3.2 *If* $R_0 > 1$, *the unique endemic equilibrium point* $P^* = (S^*, E^*, I^*)$ *of model* (3) *is globally asymptotically stable.*

Proof. The Jacobia matrix of model (3) evaluated at P^* is

$$J(P^*) = \begin{pmatrix} -\mu - \dfrac{\beta I^*}{f(I^*)} & 0 & \dfrac{(\mu+q)(\mu+\delta+\gamma)\left(I^* f'(I^*) - f(I^*)\right)}{q f(I^*)} \\ \dfrac{\beta I^*}{f(I^*)} & -\mu - q & -\dfrac{(\mu+q)(\mu+\delta+\gamma)\left(I^* f'(I^*) - f(I^*)\right)}{q f(I^*)} \\ 0 & q & -\mu - \delta - \gamma \end{pmatrix}.$$

The characteristic polynomial of $J\left(P^*\right)$ is $\lambda^3+c_1\lambda^2+c_2\lambda+c_3=0$, where

$$\begin{cases} c_1=3\mu+q+\delta+\gamma+\dfrac{\beta I^*}{f\left(I^*\right)}>0, \\ c_2=\left(\mu+\dfrac{\beta I^*}{f\left(I^*\right)}\right)\left(2\mu+q+\delta+\gamma\right)+\dfrac{(\mu+q)(\mu+\delta+\gamma)\,f'\left(I^*\right)}{f\left(I^*\right)}I^*>0, \\ c_3=\dfrac{\left(\beta+\mu f'\left(I^*\right)\right)(\mu+q)(\mu+\delta+\gamma)}{f\left(I^*\right)}I^*>0. \end{cases}$$

Clearly $c_1c_2-c_3>0$. Therefore, by the Routh-Hurwitz criterion we can conclude that P^* is locally asymptotically stable.

Next, we need to prove that P^* is globally asymptotically stable. Define a Lyapunov function

$$V\left(S,E,I\right)=\left|S-S^*\right|+\left|E-E^*\right|+\left|I-I^*\right|.$$

Obviously, $V\left(P^*\right)=0$, and when $P\neq P^*$, $V\left(P\right)\neq 0$. Remember that P^* is the solution of system (3), the upper right derivative of V can be estimated:

$$\begin{aligned} D^+V&=\operatorname{sgn}\left(S-S^*\right)\left(S-S^*\right)+\operatorname{sgn}\left(E-E^*\right)\left(E-E^*\right)+\operatorname{sgn}\left(I-I^*\right)\left(I-I^*\right) \\ &=\operatorname{sgn}\left(S-S^*\right)\left(A-\mu S-\frac{\beta IS}{f\left(I\right)}-\left(A-\mu S^*-\frac{\beta I^*S^*}{f\left(I^*\right)}\right)\right) \\ &\quad+\operatorname{sgn}\left(E-E^*\right)\left(\frac{\beta IS}{f\left(I\right)}-(\mu+q)E-\left(\frac{\beta I^*S^*}{f\left(I^*\right)}-(\mu+q)E^*\right)\right) \\ &\quad+\operatorname{sgn}\left(I-I^*\right)\left(qE-(\mu+\delta+\gamma)I-\left(qE^*-(\mu+\delta+\gamma)I^*\right)\right), \end{aligned} \tag{5}$$

where

$$\operatorname{sgn}\left(x\right)=\begin{cases} -1, & x<0, \\ 0, & x=0, \\ 1, & x>0. \end{cases}$$

In (5), there are 8 kinds of situations for the size of S and S^*, E and E^*, I and I^*. It is enough to analyze the situation of $S>S^*$, $E>E^*$ and $I>I^*$ since the other situations are similar.

Firstly there is

$$\begin{aligned} D^+V&\le-\mu\left|S-S^*\right|-\mu\left|E-E^*\right|-(\mu+\delta+\gamma)\left|I-I^*\right| \\ &<-\mu\left|S-S^*\right|-\mu\left|E-E^*\right|-\mu\left|I-I^*\right| \\ &<-\mu V. \end{aligned} \tag{6}$$

Integrating from t_0 to t on both sides of (6), we have

$$V\left(t\right)+\mu\int_{t_0}^{t}V\mathrm{d}t\le V\left(t_0\right)<+\infty.$$

Since the front set Γ has a boundary, S, E and I must have boundaries, and

their derivatives are bounded, which means that V is uniformly continuous.

By Barbalat Lemma, there is $\lim_{t\to\infty} V(t)=0$, so

$$D^{+}V<-\mu V<0.$$

Thus, the unique endemic equilibrium point P^{*} of model (3) is globally asymptotically stable. This completes the proof.

4. Applications and Numerical Simulations

In this section, we choose the function $f(I)$ as

$$f(I)=1+\alpha^{2}.$$

which was proposed by Xiao and Ruan [5].

Then the model(3) becomes

$$\begin{cases}\dfrac{\mathrm{d}S}{\mathrm{d}t}=A-\mu S-\dfrac{\beta I}{1+\alpha^{2}}S,\\ \dfrac{\mathrm{d}E}{\mathrm{d}t}=\dfrac{\beta I}{1+\alpha^{2}}S-(\mu+q)E,\\ \dfrac{\mathrm{d}I}{\mathrm{d}t}=qE-(\mu+\delta+\gamma)I.\end{cases}\tag{7}$$

It is easy to verify that the function $f(I)=1+\alpha^{2}$ satisfies assumption $H(1)$ and $H(2)$, at the same time,

$$R_{0}=\frac{\beta Aq}{\mu(\mu+q)(\mu+\delta+\gamma)}.$$

The model (7) has a disease-free equilibrium point $P_{0}=\left(\dfrac{A}{\mu},0,0\right)$ and an endemic equilibrium point $P^{*}=(S^{*},E^{*},I^{*})$ with

$$\begin{cases}S^{*}=\dfrac{\mu(\mu+q)(\mu+\delta+\gamma)}{\beta q}\left(1+\alpha^{*2}\right),\\ E^{*}=\dfrac{\mu+\delta+\gamma}{q}I^{*},\\ I^{*}=\dfrac{\sqrt{\beta^{2}+4\alpha\mu^{2}(R_{0}-1)}-\beta}{2\mu\alpha},\end{cases}$$

when $R_{0}>1$.

Using the arguments in section 3, we can obtain the following results.

Theorem 4.1 *The disease-free equilibrium point* P_{0} *of* (7) *is globally asymptotically stable if* $R_{0}\le 1$ *and unstable if* $R_{0}>1$.

Theorem 4.2 *The endemic equilibrium point* P^{*} *of* (7) *is globally asymptotically stable if* $R_{0}>1$.

We know that $\mu,\beta,q,\delta,\gamma$ are greater than 0 and less than 1, so we selected some data for numerical simulation. After many numerical simulation experiments, we have selected the following data with good interpretation.

Let $A=1$, $\mu=0.15$, $\beta=0.2$, $q=0.1$, $\delta=0.05$, $\gamma=0.1$, $\alpha=0.01$ in

Figure 1, we have $R_0 = 1.333 > 1$. As is shown in Figure 1, the curve S starts to go up very quickly, goes down at a certain value and then tends stabilize. The curves E and I go up at a certain value and flatten out. This suggests that the disease is present. Figure 1 shows that when $R_0 > 1$, the endemic equilibrium point P^* exists and is globally asymptotically stable, which is consistent with the previous conclusion.

Next, we choose some other parameter values: $A = 1$, $\mu = 0.2$, $\beta = 0.2$, $q = 0.1$, $\delta = 0.05$, $\gamma = 0.1$, $\alpha = 0.01$. We have $R_0 = 0.9523 < 1$. As shown in Figure 2, the curve S rises rapidly, reaches a certain value and then flattens out. While the curves E and I monotonically decrease and go to zero which indicate that the disease disappears over time. Figure 2 reflects that the disease dies out and the disease-free equilibrium point P_0 is globally asymptotically stable.

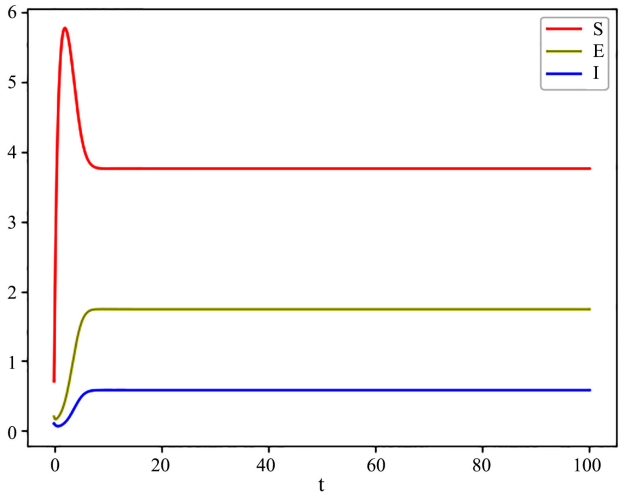


Figure 1. The path of $S(t), E(t), I(t)$ for the model (7) with initial values $(0.7, 0.2, 0.1)$, $R_0 = 1.333 > 1$.

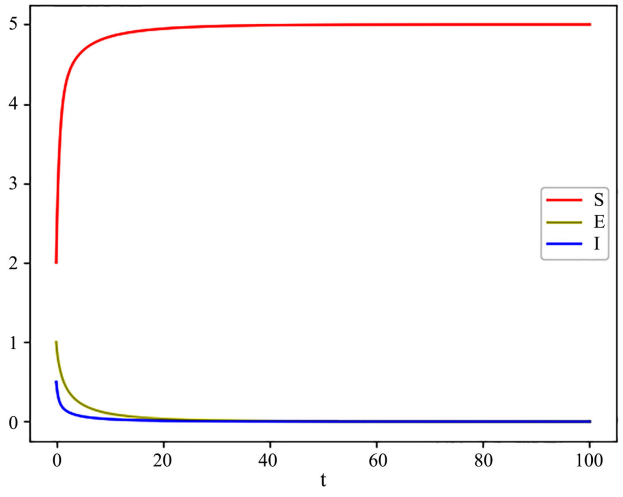


Figure 2. The path of $S(t), E(t), I(t)$ for the model (7) with initial values $(2, 1, 0.5)$, $R_0 = 0.953 < 1$.

5. Brief Summary

In this paper, we consider the global stability of an SEIR epidemic model with infection force under intervention strategies. We suppose that the infection force can be factorized into $\frac{\beta I}{f(I)}$, where $f(I)$ satisfies some conditions, and we use the regeneration matrix to obtain the basic reproductive number R_0. We also proved the existence of the equilibrium point.

We prove that if $R_0 \le 1$, there exists only the disease-free equilibrium point which is globally asymptotically stable; if $R_0 > 1$, there is a unique endemic equilibrium point and the endemic equilibrium point is globally asymptotically stable.

Conflicts of Interest

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

References

[1] Kermack, G.W.O. and Mckendrick, A.G. (1937) Contributions to the Mathematical Theory of Epidemics IV. Analysis of Experimental Epidemics of the Virus Disease Mouse Ectromelia. *Epidemiology & Infection*, **37**, 172-187. https://doi.org/10.1017/S0022172400034902

[2] Cai, Y., Kang, Y., Banerjee, M. and Wang, W. (2015) A Stochastic Sirs Epidemic Model with Infectious Force under Intervention Strategies. *Journal of Differential Equations*, **259**, 7463-7502. https://doi.org/10.1016/j.jde.2015.08.024

[3] Ma, Z., Zhou, Y. and Wu, J. (2009) Modeling and Dynamics of Infectious Diseases. Higher Education Press, Beijing.

[4] Wang, W. (2017) Epidemic Models with Nonlinear Infection Forces. *Mathematical Biosciences & Engineering*, **3**, 267-279. https://doi.org/10.3934/mbe.2006.3.267

[5] Xiao, D. and Ruan, S. (2007) Global Analysis of an Epidemic Model with Non-Monotone Incidence Rate. *Mathematical Biosciences*, **208**, 419-429. https://doi.org/10.1016/j.mbs.2006.09.025

[6] Cui, J.A., Tao, X. and Zhu, H. (2008) An Sis Infection Model Incorporating Media Coverage. *Rocky Mountain Journal of Mathematics*, **38**, 1323-1334. https://doi.org/10.1216/RMJ-2008-38-5-1323

[7] Cui, J., Sun, Y. and Zhu, H. (2008) The Impact of Media on the Control of Infectious Diseases. *Journal of Dynamics & Differential Equations*, **20**, 31-53. https://doi.org/10.1007/s10884-007-9075-0

[8] Tchuenche, J.M. and Bauch, C.T. (2012) Dynamics of an Infectious Disease Where Media Coverage Influences Transmission. *ISRN Biomathematics*, **2012**, Article ID: 581274. https://doi.org/10.5402/2012/581274

[9] Xiao, Y., Zhao, T. and Tang, S. (2013) Dynamics of an Infectious Diseases with Media/Psychology Induced Non-Smooth Incidence. *Mathematical Biosciences & Engineering*, **10**, 445-461. https://doi.org/10.3934/mbe.2013.10.445

[10] Tang, S., Xiao, Y., Yuan, L., Cheke, R.A. and Wu, J. (2012) Campus Quarantine (Fengxiao) for Curbing Emergent Infectious Diseases: Lessons from Mitigating a/h1n1 in Xi'an, China. *Journal of Theoretical Biology*, **295**, 47-58.

https://doi.org/10.1016/j.jtbi.2011.10.035

[11] Khan, M.A., Rahman, M., Khanam, P.A., Khuda, B., Kane, T.T. and Ashraf, A. (1997) Awareness of Sexually Transmitted Disease among Women and Service Providers in Rural Bangladesh. *International Journal of Std & Aids*, **8**, 688-696. https://doi.org/10.1258/0956462971919066

[12] Xiao, Y., Tang, S. and Wu, J. (2015) Media Impact Switching Surface during an Infectious Disease Outbreak. *Scientific Reports*, **5**, Article No. 7838. https://doi.org/10.1038/srep07838

[13] Ruan, S.G. and Wang, W.D. (2003) Dynamical Behavior of an Epidemic Model with a Nonlinear Incidence Rate. *Journal of Differential Equations*, **188**, 135-163. https://doi.org/10.1016/S0022-0396(02)00089-X

[14] Cai, Y., Kang, Y. and Wang, W. (2017) A Stochastic Sirs Epidemic Model with Nonlinear Incidence Rate. *Applied Mathematics & Computation*, **305**, 221-240. https://doi.org/10.1016/j.amc.2017.02.003

[15] Van den Driessche, P. and Watmough, J. (2002) Reproduction Numbers and Sub-Threshold Endemic Equilibria for Compartmental Models of Disease Transmission. *Mathematical Biosciences*, **180**, 29-48. https://doi.org/10.1016/S0025-5564(02)00108-6

[16] Bai, Z.G. (2013) Basic Reproduction Number of Periodic Epidemic Models. *Chinese Journal of Engineering Mathematics*, **30**, 175-183.

[17] LaSalle, J.P. (1976) The Stability of Dynamical Systems. CBMS-NSF Regional Conference Series in Applied Mathematics. Book 27, Society for Industrial and Applied Mathematics, Philadelphia, 1121-1130.

[18] Lyapunov, A.M. (1992) The General Problem of the Stability of Motion. *International Journal of Control*, **55**, 531-534. https://doi.org/10.1080/00207179208934253

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1718-1731
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Algebraic Techniques for Linear Equations over Quaternions and Split Quaternions: A Unified Approach in Quaternionic and Split Quaternionic Mechanics

Gang Wang[1,2], Zhenwei Guo[1], Dong Zhang[3], Tongsong Jiang[1,2,3*]

[1]School of Mathematical Science, Liaocheng University, Liaocheng, China
[2]School of Mathematics and Statistics, Heze University, Heze, China
[3]College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao, China
Email: wang_gang93@163.com, *jiangtongsong@sina.com, dz_zhangdong@sina.com

How to cite this paper: Wang, G., Guo, Z.W., Zhang, D. and Jiang, T.S. (2019) Algebraic Techniques for Linear Equations over Quaternions and Split Quaternions: A Unified Approach in Quaternionic and Split Quaternionic Mechanics. *Journal of Applied Mathematics and Physics*, **7**, 1718-1731.
https://doi.org/10.4236/jamp.2019.78118

Received: July 18, 2019
Accepted: August 13, 2019
Published: August 16, 2019

Abstract

This paper aims to present, in a unified manner, algebraic techniques for linear equations which are valid on both the algebras of quaternions and split quaternions. This paper, introduces a concept of v-quaternion, studies the problem of v-quaternionic linear equations by means of a complex representation and a real representation of v-quaternion matrices, and gives two algebraic methods for solving v-quaternionic linear equations. This paper also gives a unification of algebraic techniques for quaternionic and split quaternionic linear equations in quaternionic and split quaternionic mechanics.

Keywords

V-Quaternion, Complex Representation, Real Representation, Linear Equations, Quaternion, Split Quaternion

1. Introduction

A quaternion, which was found in 1840 by William Rowan Hamilton [1], is in the form of $q = q_1 + q_2\mathrm{i} + q_3\mathrm{j} + q_4\mathrm{k}$, $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$, $\mathrm{ijk} = -1$, where $q_1, q_2, q_3, q_4 \in \mathbf{R}$, and $\mathrm{ij} = -\mathrm{ji} = \mathrm{k}$, $\mathrm{jk} = -\mathrm{kj} = \mathrm{i}$, $\mathrm{ki} = -\mathrm{ik} = \mathrm{j}$. Quaternion algebra has been playing a significant role recently in geometric and physical applications, many geometric problems can be represented by quaternions. In paper [2], the authors showed that a unit timelike quaternion represents a rotation in the Minkowski 3 space, and expressed Lorentzian rotation matrix generated

with a time like quaternion. In paper [3], the authors studied the problem of using quaternions in unconstrained nonlinear optimization of 3-D rotations, and gave an easy and accurate method for applying the quaternion representation of 3-D rotations.

A split quaternion (or coquaternion), which was found in 1849 by James Cockle [4], is in the form of $q = q_1 + q_2\mathrm{i} + q_3\mathrm{j} + q_4\mathrm{k}$, $\mathrm{i}^2 = -1$, $\mathrm{j}^2 = \mathrm{k}^2 = 1$, $\mathrm{ijk} = 1$, where $q_1, q_2, q_3, q_4 \in \mathbf{R}$, and $\mathrm{ij} = -\mathrm{ji} = \mathrm{k}$, $\mathrm{jk} = -\mathrm{kj} = -\mathrm{i}$, $\mathrm{ki} = -\mathrm{ik} = \mathrm{j}$ and denotes the sets of quaternions and split quaternions respectively by $\mathbf{H}$ and $\mathbf{H}_s$. The quaternion ring $\mathbf{H}$ and the split quaternion ring $\mathbf{H}_s$ are two associative and noncommutative 4-dimensional Clifford algebras, and the split quaternion ring $\mathbf{H}_s$ contains zero divisors, nilpotent elements and nontrivial idempotents. In paper [5], the authors stated the rotations in Minkowski 3 space by split quaternions. In paper [6], the authors studied dual split quaternions and screw motion in 3-dimensional Lorentzian space, and obtained the components of a dual split quaternion by replacing the L-Euler parameters with their split dual versions. In paper [7], the authors studied eigenvalue problem of a rotation matrix in Minkowski 3 space by using split quaternions, and gave the characterizations of eigenvalues of a rotation matrix in Minkowski 3 space according to only first component of the corresponding quaternion. Quaternions and split quaternions in the study of geometry and physic are more than those, e.g. [8]-[13].

A v-quaternion is in the form of

$$q = q_1 + q_2\mathrm{i} + q_3\mathrm{j} + q_4\mathrm{k},\ \mathrm{i}^2 = -1,\ \mathrm{j}^2 = v,\ \mathrm{ij} = -\mathrm{ji} = \mathrm{k}, \tag{1.1}$$

in which $0 \neq v \in \mathbf{R}$, $q_1, q_2, q_3, q_4 \in \mathbf{R}$, and $\mathrm{k}^2 = \mathrm{ijk} = v$, $\mathrm{jk} = -\mathrm{kj} = -v\mathrm{i}$, $\mathrm{ik} = -\mathrm{ki} = -\mathrm{j}$. Let $\mathbf{H}_v$ denote the set of v-quaternion. Obviously, the set of all v-quaternion is also a noncommutative 4-dimensional Clifford algebra. Specially, when $v = -1$, the ring of the v-quaternion $\mathbf{H}_v$ is the ring of the quaternion $\mathbf{H}$; when $v = 1$, the ring of the v-quaternion $\mathbf{H}_v$ is the ring of the split quaternion $\mathbf{H}_s$.

In the geometry research and physical application of quaternion and split quaternion, the problems of solving quaternionic and split quaternionic equations are often encountered. In paper [14], by means of a complex representation of quaternion matrices, the authors studied the problems of quaternionic linear equations, and gave an algorithm for quaternionic linear equations. In paper [15], by means of a complex representation and a real representation of split quaternion matrices, the authors studied the split quaternionic least squares problem, and derived two algebraic methods for finding solutions of the problems in split quaternionic mechanics. For the problems of quaternions and split quaternions, the scholars need to discuss by classification. However, as two special cases of four-dimensional algebra, it is of theoretical and practical significance to solve them in a unified way. This paper aims to present, in a unified manner, algebraic techniques for linear equations which are valid on both the algebras of quaternions and split quaternions. This paper, by means of a complex represen-

tation and a real representation of v-quaternion matrices, studies the problem of v-quaternionic linear equations, and gives two algebraic methods for solving v-quaternionic linear equations. This paper also gives a unification of algebraic techniques for quaternionic and split quaternionic linear equations in quaternionic and split quaternionic mechanics.

Let $\mathbf{R}$ be the real number field, $\mathbf{C}=\mathbf{R}\oplus\mathbf{R}\mathrm{i}$ the complex number field. If $q=q_1+q_2\mathrm{i}+q_3\mathrm{j}+q_4\mathrm{k}\in\mathbf{H}_v$, $\overline{q}=q_1-q_2\mathrm{i}-q_3\mathrm{j}-q_4\mathrm{k}$ is the conjugate of q. For any matrix $A=(a_{st})\in\mathbf{H}_v^{m\times n}$, $\overline{A}=(\overline{a}_{st})$, $A^{\mathrm{T}}=(a_{ts})$, $A^{*}=(\overline{a}_{ts})$, A^{-1} denote the conjugate, the transpose, the conjugate transpose and the inverse of the matrix A, respectively.

This paper is organized as follows. In Section 2, we give two new matrix representations of v-quaternion matrix, and discuss some properties and conclusions of complex representation and real representation of v-quaternion matrices. In Section 3, we present the complex representation method for solving v-quaternionic linear equations and some numerical examples. In Section 4, we present the real representation method for solving v-quaternionic linear equations and some numerical examples. In Section 5, we summarize this paper.

2. Complex Representation and Real Representation of V-Quaternion Matrices

For any v-quaternion matrix
$A=A_1+A_2\mathrm{i}+A_3\mathrm{j}+A_4\mathrm{k}=(A_1+A_2\mathrm{i})+(A_3+A_4\mathrm{i})\mathrm{j}=B_1+B_2\mathrm{j}\in\mathbf{H}_v^{m\times n}$,
$A_1,A_2,A_3,A_4\in\mathbf{R}^{m\times n}$, $B_1,B_2\in\mathbf{C}^{m\times n}$, the complex representation A^C of the v-quaternion matrix A is defined to be

$$A^C=\begin{bmatrix}B_1 & vB_2\\ \overline{B}_2 & \overline{B}_1\end{bmatrix}, \tag{2.1}$$

and the real representation A^R of the v-quaternion matrix A is defined to be

$$A^R=\begin{bmatrix}A_1 & -A_2 & vA_3 & vA_4\\ A_2 & A_1 & vA_4 & -vA_3\\ A_3 & A_4 & A_1 & -A_2\\ A_4 & -A_3 & A_2 & A_1\end{bmatrix}. \tag{2.2}$$

For any v-quaternion matrix $A,B\in\mathbf{H}_v^{m\times n}$, $C\in\mathbf{H}_v^{n\times p}$, $a\in\mathbf{R}$, for $\sigma\in\{C,R\}$, it is easy to prove the following equalities by direct calculation.

$$(A+B)^{\sigma}=A^{\sigma}+B^{\sigma},(aA)^{\sigma}=aA^{\sigma},(AC)^{\sigma}=A^{\sigma}C^{\sigma}, \tag{2.3}$$

and

$$Q_m^{-1}A^CQ_n=\overline{A^C}, \tag{2.4}$$

where $Q_t=\begin{bmatrix}0 & vI_t\\ I_t & 0\end{bmatrix}$.

Similarly, by direct calculation we get the following results.

$$P_m^{-1}A^RP_n=A^R,R_m^{-1}A^RR_n=A^R,S_m^{-1}A^RS_n=A^R, \tag{2.5}$$

where $P_t = \begin{bmatrix} 0 & -I_t & 0 & 0 \\ I_t & 0 & 0 & 0 \\ 0 & 0 & 0 & I_t \\ 0 & 0 & -I_t & 0 \end{bmatrix}$, $R_t = \begin{bmatrix} 0 & 0 & vI_t & 0 \\ 0 & 0 & 0 & vI_t \\ I_t & 0 & 0 & 0 \\ 0 & I_t & 0 & 0 \end{bmatrix}$,

$S_t = \begin{bmatrix} 0 & 0 & 0 & vI_t \\ 0 & 0 & -vI_t & 0 \\ 0 & -I_t & 0 & 0 \\ I_t & 0 & 0 & 0 \end{bmatrix}$, and $P_t^{-1} = -P_t$, $R_t^{-1} = \frac{1}{v}R_t$, $S_t^{-1} = \frac{1}{v}S_t$.

Lemma 2.1 For two special cases of quaternion $(v=-1)$ and split quaternion $(v=1)$ matrices, clearly by (2.1) and (2.2) the complex representation and the real representation are respectively to be

$$A^C = \begin{bmatrix} B_1 & -B_2 \\ \overline{B}_2 & \overline{B}_1 \end{bmatrix}, A^R = \begin{bmatrix} A_1 & -A_2 & -A_3 & -A_4 \\ A_2 & A_1 & -A_4 & A_3 \\ A_3 & A_4 & A_1 & -A_2 \\ A_4 & -A_3 & A_2 & A_1 \end{bmatrix}, A \in \mathbf{H}^{m\times n}. \tag{2.6a}$$

$$A^C = \begin{bmatrix} B_1 & B_2 \\ \overline{B}_2 & \overline{B}_1 \end{bmatrix}, A^R = \begin{bmatrix} A_1 & -A_2 & A_3 & A_4 \\ A_2 & A_1 & A_4 & -A_3 \\ A_3 & A_4 & A_1 & -A_2 \\ A_4 & -A_3 & A_2 & A_1 \end{bmatrix}, A \in \mathbf{H}_s^{m\times n}. \tag{2.6b}$$

For any v-quaternion matrix $A \in \mathbf{H}_v^{m\times n}$, the rank $\operatorname{rank}(A)$ of the matrix A is defined to be

$$\operatorname{rank}(A) \equiv \frac{1}{2}\operatorname{rank}\left(A^C\right), \tag{2.7}$$

or

$$\operatorname{rank}(A) \equiv \frac{1}{4}\operatorname{rank}\left(A^R\right). \tag{2.8}$$

By the definition of rank and (2.3), it is easy to get the following results by direct calculation. If $A, B \in \mathbf{H}_v^{m\times n}$, $C \in \mathbf{H}_v^{n\times p}$, then $\operatorname{rank}(A+B) \le \operatorname{rank}(A) + \operatorname{rank}(B)$ and $\operatorname{rank}(AC) \le \min\{\operatorname{rank}(A), \operatorname{rank}(C)\}$.

3. Algebraic Method of Complex Representation

If $A \in \mathbf{H}_v^{m\times n}, B \in \mathbf{H}_v^{m\times p}$, then by the definition of complex representation and (2.3), $AX = B$ if and only if $A^C X^C = B^C$. That is $AX = B$ has a solution X if and only if $A^C Y = B^C$ has a solution $Y = X^C$.

Theorem 3.1 For $A \in \mathbf{H}_v^{m\times n}, B \in \mathbf{H}_v^{m\times p}$. Then

1) V-quaternionic linear equations $AX = B$ have a solution if and only if $\operatorname{rank}(A) = \operatorname{rank}(A, B)$, *i.e.* $AX = B$ has a solution if and only if $A^C Y = B^C$ has a solution, and if $\operatorname{rank}(A) = \operatorname{rank}(A, B) = n$, then v-quaternionic linear equations $AX = B$ have a unique solution.

2) If Y is a solution to $A^C Y = B^C$, then the following v-quaternion matrix is a solution to $AX = B$,

$$X = \frac{1}{4}\left[I_n, I_n \mathrm{j}\right]\left(Y + Q_n^{-1}\overline{Y}Q_p\right)\begin{bmatrix} I_p \\ \frac{1}{v}I_p \mathrm{j}\end{bmatrix}, \tag{3.1}$$

in which $Q_t = \begin{bmatrix} 0 & vI_t \\ I_t & 0 \end{bmatrix}$.

Proof: If Y is a solution of $A^C Y = B^C$, by (2.4),

$$A^C Y = B^C \Leftrightarrow \overline{A^C}\left(Q_n^{-1} Y Q_p\right) = \overline{B^C} \Leftrightarrow A^C\left(Q_n^{-1}\overline{Y}Q_p\right) = B^C, \tag{3.2}$$

i.e. $Q_n^{-1}\overline{Y}Q_p$ is a solution of $A^C Y = B^C$, therefore

$$\hat{Y} = \frac{1}{2}\left(Y + Q_n^{-1}\overline{Y}Q_p\right) \tag{3.3}$$

is also a solution of $A^C Y = B^C$. Let

$$Y = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \in \mathbf{C}^{2n\times 2p},\ z_{ts} \in \mathbf{C}^{n\times p},\ s,t = 1,2. \tag{3.4}$$

It is easy to get, by direct calculation,

$$\hat{Y} = \begin{bmatrix} \hat{z}_1 & v\hat{z}_2 \\ \overline{\hat{z}}_2 & \overline{\hat{z}}_1 \end{bmatrix} \in \mathbf{C}^{2n\times 2p}, \tag{3.5}$$

in which

$$\hat{z}_1 = \frac{1}{2}\left(z_{11} + \overline{z}_{22}\right),\ \hat{z}_2 = \frac{1}{2}\left(\frac{1}{v}z_{12} + \overline{z}_{21}\right). \tag{3.6}$$

By (3.5), we construct a v-quaternion matrix.

$$X = \hat{z}_1 + \hat{z}_2 \mathrm{j} = \frac{1}{2}\left[I_n, I_n \mathrm{j}\right]\hat{Y}\begin{bmatrix} I_p \\ \frac{1}{v}I_p \mathrm{j}\end{bmatrix}. \tag{3.7}$$

Clearly $X^C = \hat{Y}$. This means that $X^C = \hat{Y}$ is a solution of $A^C Y = B^C$, so X is a solution of $AX = B$.

From the statement above we get following results. When the v-quaternionic linear equations $AX = B$ have a solution, we can find a solution by a solution of complex representation equation $A^C Y = B^C$ from the formula (3.1).

The following two special cases about quaternions and split quaternions come from Theorem 3.1 respectively with $v = -1$ and $v = 1$.

Corollary 3.2 For $A \in \mathbf{H}^{m\times n}, B \in \mathbf{H}^{m\times p}$. Then

1) The quaternionic linear equations $AX = B$ have a solution if and only if $\mathrm{rank}(A) = \mathrm{rank}(A, B)$, *i.e.* $AX = B$ has a solution if and only if $A^C Y = B^C$ has a solution, and if $\mathrm{rank}(A) = \mathrm{rank}(A, B) = n$, then quaternionic linear equations $AX = B$ have a unique solution.

2) If Y is a solution to $A^C Y = B^C$, then the following quaternion matrix is a solution to $AX = B$,

$$X = \frac{1}{4}\left[I_n, I_n \mathrm{j}\right]\left(Y + Q_n^{-1}\overline{Y}Q_p\right)\begin{bmatrix} I_p \\ -I_p \mathrm{j}\end{bmatrix}, \tag{3.8}$$

in which $Q_t = \begin{bmatrix} 0 & -I_t \\ I_t & 0 \end{bmatrix}$.

Corollary 3.3 *For* $A \in \mathbf{H}_s^{m\times n}, B \in \mathbf{H}_s^{m\times p}$. *Then*

1) The split quaternionic linear equations $AX = B$ have a solution if and only if $\operatorname{rank}(A) = \operatorname{rank}(A, B)$, *i.e.* $AX = B$ has a solution if and only if $A^C Y = B^C$ has a solution, and if $\operatorname{rank}(A) = \operatorname{rank}(A, B) = n$, then split quaternionic linear equations $AX = B$ have a unique solution.

2) If Y is a solution to $A^C Y = B^C$, then the following split quaternion matrix is a solution to $AX = B$,

$$X = \frac{1}{4}[I_n, I_n \mathrm{j}]\left(Y + Q_n^{-1}\overline{Y}Q_p\right)\begin{bmatrix} I_p \\ I_p \mathrm{j} \end{bmatrix}, \tag{3.9}$$

in which $Q_t = \begin{bmatrix} 0 & I_t \\ I_t & 0 \end{bmatrix}$.

In the similarly way, we have the following result.

Theorem 3.4 For $A \in \mathbf{H}_v^{m\times n}, C \in \mathbf{H}_v^{p\times q}, B \in \mathbf{H}_v^{m\times q}$. Then

1) V-quaternionic matrix equation $AXC = B$ has a solution if and only if $\operatorname{rank}(A) = \operatorname{rank}(A, B)$ and $\operatorname{rank}(C) = \operatorname{rank}\begin{pmatrix} C \\ B \end{pmatrix}$, *i.e.* $AXC = B$ has a solution if and only if $A^C Y C^C = B^C$ has a solution, and if $\operatorname{rank}(A) = \operatorname{rank}(A, B) = n$, $\operatorname{rank}(C) = \operatorname{rank}\begin{pmatrix} C \\ B \end{pmatrix} = p$, then v-quaternionic matrix equation $AXC = B$ has a unique solution.

2) If Y is a solution to $A^C Y C^C = B^C$, then the following v-quaternion matrix is a solution to $AXC = B$,

$$X = \frac{1}{4}[I_n, I_n \mathrm{j}]\left(Y + Q_n^{-1}\overline{Y}Q_p\right)\begin{bmatrix} I_p \\ \frac{1}{v} I_p \mathrm{j} \end{bmatrix}, \tag{3.10}$$

in which $Q_t = \begin{bmatrix} 0 & vI_t \\ I_t & 0 \end{bmatrix}$.

The proof process is similar to the Theorem 3.1.

Remark 1 The above theorems and corollaries not only give the necessary and sufficient conditions for quaternion and split quaternion matrix equations $AX = B, AXC = B$ to have a solution, but also a unification of representation for a solution.

Example 3.1

Let

$$A = \begin{pmatrix} \mathrm{i} & 1+\mathrm{j} \\ -1+\mathrm{j} & -\mathrm{k} \end{pmatrix} \text{ and } B = \begin{pmatrix} \mathrm{i} \\ -1 \end{pmatrix}.$$

Find all solutions of the v-quaternionic linear equations $AX = B$.

By the complex representation of the v-quaternion matrix, we know

$$A^C=\begin{bmatrix} \mathrm{i} & 1 & 0 & v \\ -1 & 0 & v & -v\mathrm{i} \\ 0 & 1 & -\mathrm{i} & 1 \\ 1 & \mathrm{i} & -1 & 0 \end{bmatrix}, B^C=\begin{bmatrix} \mathrm{i} & 0 \\ -1 & 0 \\ 0 & -\mathrm{i} \\ 0 & -1 \end{bmatrix},$$

and if $v\neq\frac{7\pm\sqrt{45}}{2}$, then $\operatorname{rank}\left(A^C\right)=\operatorname{rank}\left(A^C,B^C\right)=4$, *i.e.* $\operatorname{rank}(A)=\operatorname{rank}(A,B)=2$, then the v-quaternionic linear equations $AX=B$ have a unique solution.

For the matrix equation $A^CY=B^C$, the unique solution is easily found to be

$$Y=\begin{bmatrix} \dfrac{-5v+1}{v^2-7v+1} & \dfrac{v(1-2v)}{v^2-7v+1} \\ \dfrac{-3v}{v^2-7v+1}\mathrm{i} & \dfrac{v(-1-v)}{v^2-7v+1}\mathrm{i} \\ \dfrac{1-2v}{v^2-7v+1} & \dfrac{-5v+1}{v^2-7v+1} \\ \dfrac{1+v}{v^2-7v+1}\mathrm{i} & \dfrac{3v}{v^2-7v+1}\mathrm{i} \end{bmatrix}.$$

By (3.1), we easily find the unique solution X of v-quaternionic linear equations $AX=B$, and

$$\begin{aligned} X&=\frac{1}{4}\left[I_2,I_2\mathrm{j}\right]\left(Y+Q_n^{-1}\overline{Y}Q_p\right)\begin{bmatrix} 1 \\ \dfrac{1}{v}\mathrm{j} \end{bmatrix} \\ &=\left[\frac{-5v+1}{v^2-7v+1}+\frac{1-2v}{v^2-7v+1}\mathrm{j}\quad \frac{-3v}{v^2-7v+1}\mathrm{i}-\frac{1+v}{v^2-7v+1}\mathrm{k}\right]^{\mathrm{T}}. \end{aligned}$$

The following two examples are special cases of the above conclusion.

Case 1: For quaternionic linear equations $AX=B$ with $v=-1$. It is easy to know A^C and B^C by (2.6a),

$$A^C=\begin{bmatrix} \mathrm{i} & 1 & 0 & -1 \\ -1 & 0 & -1 & \mathrm{i} \\ 0 & 1 & -\mathrm{i} & 1 \\ 1 & \mathrm{i} & -1 & 0 \end{bmatrix}, B^C=\begin{bmatrix} \mathrm{i} & 0 \\ -1 & 0 \\ 0 & -\mathrm{i} \\ 0 & -1 \end{bmatrix}$$

and $\operatorname{rank}\left(A^C\right)=\operatorname{rank}\left(A^C,B^C\right)=4$. Clearly, the linear equations $A^CY=B^C$ have a unique solution. The unique solution is easily found to be

$$Y=\begin{bmatrix} \dfrac{2}{3} & -\dfrac{1}{3} \\ \dfrac{1}{3}\mathrm{i} & 0 \\ \dfrac{1}{3} & \dfrac{2}{3} \\ 0 & -\dfrac{1}{3}\mathrm{i} \end{bmatrix}.$$

By (3.8), we easily find the unique solution X of quaternionic linear equations

$AX = B$, and

$$X = \frac{1}{4}[I_2, I_2\mathrm{j}]\left(Y + Q_n^{-1}\overline{Y}Q_p\right)\begin{bmatrix}1\\-\mathrm{j}\end{bmatrix} = \begin{bmatrix}\frac{2}{3}+\frac{1}{3}\mathrm{j} & \frac{1}{3}\mathrm{i}\end{bmatrix}^{\mathrm{T}}.$$

Case 2: For split quaternionic linear equations $AX = B$ with $v = 1$. It is easy to know A^C and B^C by (2.6b),

$$A^C = \begin{bmatrix}\mathrm{i} & 1 & 0 & 1\\ -1 & 0 & 1 & -\mathrm{i}\\ 0 & 1 & -\mathrm{i} & 1\\ 1 & \mathrm{i} & -1 & 0\end{bmatrix}, B^C = \begin{bmatrix}\mathrm{i} & 0\\ -1 & 0\\ 0 & -\mathrm{i}\\ 0 & -1\end{bmatrix}$$

and $\mathrm{rank}\left(A^C\right) = \mathrm{rank}\left(A^C, B^C\right) = 4$. Clearly, the linear equations $A^C Y = B^C$ have a unique solution. The unique solution is easily found to be

$$Y = \begin{bmatrix}\frac{4}{5} & \frac{1}{5}\\ \frac{3}{5}\mathrm{i} & \frac{2}{5}\mathrm{i}\\ \frac{1}{5} & \frac{4}{5}\\ -\frac{2}{5}\mathrm{i} & -\frac{3}{5}\mathrm{i}\end{bmatrix}.$$

By (3.9), we easily find the unique solution X of split quaternionic linear equations $AX = B$, and

$$X = \frac{1}{4}[I_2, I_2\mathrm{j}]\left(Y + Q_n^{-1}\overline{Y}Q_p\right)\begin{bmatrix}1\\\mathrm{j}\end{bmatrix} = \begin{bmatrix}\frac{4}{5}+\frac{1}{5}\mathrm{j} & \frac{3}{5}\mathrm{i}+\frac{2}{5}\mathrm{k}\end{bmatrix}^{\mathrm{T}}.$$

4. Algebraic Method of Real Representation

If $A \in \mathbf{H}_v^{m\times n}, B \in \mathbf{H}_v^{m\times p}$, then by the definition of real representation, $AX = B$ if and only if $A^R X^R = B^R$. That is $AX = B$ has a solution X if and only if $A^R Y = B^R$ has a solution $Y = X^R$.

Theorem 4.1 For $A \in \mathbf{H}_v^{m\times n}, B \in \mathbf{H}_v^{m\times p}$. Then

1) V-quaternionic linear equations $AX = B$ have a solution if and only if $\mathrm{rank}(A) = \mathrm{rank}(A, B)$, *i.e.* $AX = B$ has a solution if and only if $A^R Y = B^R$ has a solution, and if $\mathrm{rank}(A) = \mathrm{rank}(A, B) = n$, then v-quaternionic linear equations $AX = B$ have a unique solution.

2) If Y is a solution to $A^R Y = B^R$, then the following v-quaternion matrix is a solution to $AX = B$,

$$X = \frac{1}{16}[I_n, I_n\mathrm{i}, I_n\mathrm{j}, I_n\mathrm{k}]\left(Y - P_n Y P_p + \frac{1}{v}R_n Y R_p + \frac{1}{v}S_n Y S_p\right)\begin{bmatrix}I_p\\ -I_p\mathrm{i}\\ \frac{1}{v}I_p\mathrm{j}\\ \frac{1}{v}I_p\mathrm{k}\end{bmatrix}. \quad (4.1)$$

Proof: If Y is a solution of $A^R Y = B^R$, by (2.5),

$$A^R Y = B^R \Leftrightarrow P_m^{-1} A^R P_n \left(P_n^{-1} Y P_p \right) = P_m^{-1} B^R P_p \Leftrightarrow A^R \left(P_n^{-1} Y P_p \right) = B^R, \quad (4.2)$$

i.e. $P_n^{-1} Y P_p$ is a solution of $A^R Y = B^R$. Similarly, $R_n^{-1} Y R_p, S_n^{-1} Y S_p$ are also solution of $A^R Y = B^R$.

$$\begin{aligned} \hat{Y} &= \frac{1}{4}\left(Y + P_n^{-1} Y P_p + R_n^{-1} Y R_p + S_n^{-1} Y S_p \right) \\ &= \frac{1}{4}\left(Y - P_n Y P_p + \frac{1}{v} R_n Y R_p + \frac{1}{v} S_n Y S_p \right) \end{aligned} \quad (4.3)$$

is also a solution of $A^R Y = B^R$. Let

$$Y = \begin{bmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{21} & z_{22} & z_{23} & z_{24} \\ z_{31} & z_{32} & z_{33} & z_{34} \\ z_{41} & z_{42} & z_{43} & z_{44} \end{bmatrix} \in \mathbf{R}^{4n \times 4p},\ z_{ts} \in \mathbf{R}^{n \times p},\ s,t = 1,2,3,4. \quad (4.4)$$

It is easy to get, by direct calculation,

$$\hat{Y} = \begin{bmatrix} \hat{z}_1 & -\hat{z}_2 & v\hat{z}_3 & v\hat{z}_4 \\ \hat{z}_2 & \hat{z}_1 & v\hat{z}_4 & -v\hat{z}_3 \\ \hat{z}_3 & \hat{z}_4 & \hat{z}_1 & -\hat{z}_2 \\ \hat{z}_4 & -\hat{z}_3 & \hat{z}_2 & \hat{z}_1 \end{bmatrix} \in \mathbf{R}^{4n \times 4p}, \quad (4.5)$$

in which

$$\hat{z}_1 = \frac{1}{4}\left(z_{11} + z_{22} + z_{33} + z_{44} \right), \hat{z}_2 = \frac{1}{4}\left(z_{21} - z_{12} + z_{43} - z_{34} \right), \quad (4.6a)$$

$$\hat{z}_3 = \frac{1}{4}\left(z_{31} - z_{42} + \frac{1}{v} z_{13} - \frac{1}{v} z_{24} \right), \hat{z}_4 = \frac{1}{4}\left(z_{41} + z_{32} + \frac{1}{v} z_{23} + \frac{1}{v} z_{14} \right). \quad (4.6b)$$

By (4.5), we construct a v-quaternion matrix.

$$X = \hat{z}_1 + \hat{z}_2 \mathrm{i} + \hat{z}_3 \mathrm{j} + \hat{z}_4 \mathrm{k} = \frac{1}{4}\left[I_n, I_n \mathrm{i}, I_n \mathrm{j}, I_n \mathrm{k} \right] \hat{Y} \begin{bmatrix} I_p \\ -I_p \mathrm{i} \\ \frac{1}{v} I_p \mathrm{j} \\ \frac{1}{v} I_p \mathrm{k} \end{bmatrix}. \quad (4.7)$$

Clearly $X^R = \hat{Y}$. This means that $X^R = \hat{Y}$ is a solution of $A^R Y = B^R$, so X is a solution of $AX = B$.

From the statement above we get following results.

The following two special cases about quaternions and split quaternions come from Theorem 4.1 respectively with $v = -1$ and $v = 1$.

Corollary 4.2 For $A \in \mathbf{H}^{m \times n}, B \in \mathbf{H}^{m \times p}$. Then

1) The quaternionic linear equations $AX = B$ have a solution if and only if $\mathrm{rank}(A) = \mathrm{rank}(A, B)$, *i.e.* $AX = B$ has a solution if and only if $A^R Y = B^R$ has a solution, and if $\mathrm{rank}(A) = \mathrm{rank}(A, B) = n$, then quaternionic linear equations $AX = B$ have a unique solution.

2) If Y is a solution to $A^R Y = B^R$, then the following quaternion matrix is a solution to $AX = B$,

$$X=\frac{1}{16}\left[I_n,I_n\mathrm{i},I_n\mathrm{j},I_n\mathrm{k}\right]\left(Y-P_nYP_p-R_nYR_p-S_nYS_p\right)\begin{bmatrix}I_p\\-I_p\mathrm{i}\\-I_p\mathrm{j}\\-I_p\mathrm{k}\end{bmatrix}. \tag{4.8}$$

Corollary 4.3 For $A\in\mathbf{H}_s^{m\times n}, B\in\mathbf{H}_s^{m\times p}$. Then

1) The split quaternionic linear equations $AX=B$ have a solution if and only if $\mathrm{rank}(A)=\mathrm{rank}(A,B)$, *i.e.* $AX=B$ has a solution if and only if $A^RY=B^R$ has a solution, and if $\mathrm{rank}(A)=\mathrm{rank}(A,B)=n$, then split quaternionic linear equations $AX=B$ have a unique solution.

2) If Y is a solution to $A^RY=B^R$, then the following split quaternion matrix is a solution to $AX=B$,

$$X=\frac{1}{16}\left[I_n,I_n\mathrm{i},I_n\mathrm{j},I_n\mathrm{k}\right]\left(Y-P_nYP_p+R_nYR_p+S_nYS_p\right)\begin{bmatrix}I_p\\-I_p\mathrm{i}\\I_p\mathrm{j}\\I_p\mathrm{k}\end{bmatrix}. \tag{4.9}$$

In the similarly way, we have the following result.

Theorem 4.4 For $A\in\mathbf{H}_v^{m\times n}, C\in\mathbf{H}_v^{p\times q}, B\in\mathbf{H}_v^{m\times q}$. Then

1) V-quaternionic matrix equation $AXC=B$ has a solution if and only if $\mathrm{rank}(A)=\mathrm{rank}(A,B)$ and $\mathrm{rank}(C)=\mathrm{rank}\begin{pmatrix}C\\B\end{pmatrix}$, *i.e.* $AXC=B$ has a solution if and only if $A^RYC^R=B^R$ has a solution, and if $\mathrm{rank}(A)=\mathrm{rank}(A,B)=n$, $\mathrm{rank}(C)=\mathrm{rank}\begin{pmatrix}C\\B\end{pmatrix}=p$, then v-quaternionic matrix equation $AXC=B$ has a unique solution.

2) If Y is a solution to $A^RYC^R=B^R$, then the following v-quaternion matrix is a solution to $AXC=B$,

$$X=\frac{1}{16}\left[I_n,I_n\mathrm{i},I_n\mathrm{j},I_n\mathrm{k}\right]\left(Y-P_nYP_p+\frac{1}{v}R_nYR_p+\frac{1}{v}S_nYS_p\right)\begin{bmatrix}I_p\\-I_p\mathrm{i}\\\frac{1}{v}I_p\mathrm{j}\\\frac{1}{v}I_p\mathrm{k}\end{bmatrix}. \tag{4.10}$$

The proof process is similar to the Theorem 4.1.

Remark 2 The above theorems and corollaries not only give the necessary and sufficient conditions for quaternion and split quaternion matrix equations $AX=B, AXC=B$ to have a solution, but also a unification of representation for a solution.

Example 4.1

For two v-quaternion matrices A and B in Example 3.1, find solutions of the v-quaternionic linear equations $AX=B$.

By the real representation of the v-quaternion matrix, we know

$$A^R=\begin{bmatrix}0&1&-1&0&0&v&0&0\\-1&0&0&0&v&0&0&-v\\1&0&0&1&0&0&0&-v\\0&0&-1&0&0&-v&-v&0\\0&1&0&0&0&1&-1&0\\1&0&0&-1&-1&0&0&0\\0&0&0&-1&1&0&0&1\\0&-1&-1&0&0&0&-1&0\end{bmatrix},B^R=\begin{bmatrix}0&-1&0&0\\-1&0&0&0\\1&0&0&0\\0&-1&0&0\\0&0&0&-1\\0&0&-1&0\\0&0&1&0\\0&0&0&-1\end{bmatrix},$$

and if $v\neq\frac{7\pm\sqrt{45}}{2}$, then $\operatorname{rank}\left(A^R\right)=\operatorname{rank}\left(A^R,B^R\right)=8$, *i.e.* $\operatorname{rank}\left(A\right)=\operatorname{rank}\left(A,B\right)=2$, then the v-quaternionic linear equations $AX=B$ have a unique solution.

For the matrix equation $A^RY=B^R$, the unique solution is easily found to be

$$Y=\frac{1}{v^2-7v+1}\begin{bmatrix}-5v+1&0&(1-2v)v&0\\0&3v&0&(-1-v)v\\0&-5v+1&0&(2v-1)v\\-3v&0&(-1-v)v&0\\1-2v&0&-5v+1&0\\0&-1-v&0&3v\\0&2v-1&0&-5v+1\\-1-v&0&-3v&0\end{bmatrix}.$$

By (4.1), we easily find the unique solution X of v-quaternionic linear equations $AX=B$, and

$$X=\frac{1}{16}\left[I_2,I_2\mathrm{i},I_2\mathrm{j},I_2\mathrm{k}\right]\left(Y-P_nYP_p+\frac{1}{v}R_nYR_p+\frac{1}{v}S_nYS_p\right)\begin{bmatrix}1\\-\mathrm{i}\\\frac{1}{v}\mathrm{j}\\\frac{1}{v}\mathrm{k}\end{bmatrix}$$

$$=\left[\frac{-5v+1}{v^2-7v+1}+\frac{1-2v}{v^2-7v+1}\mathrm{j}\quad\frac{-3v}{v^2-7v+1}\mathrm{i}-\frac{1+v}{v^2-7v+1}\mathrm{k}\right]^{\mathrm{T}}.$$

The following two examples are special cases of the above conclusion.

Case 1: For quaternionic linear equations $AX=B$ with $v=-1$. It is easy to know A^R and B^R by (2.6a),

$$A^R=\begin{bmatrix}0&1&-1&0&0&-1&0&0\\-1&0&0&0&-1&0&0&1\\1&0&0&1&0&0&0&1\\0&0&-1&0&0&1&1&0\\0&1&0&0&0&1&-1&0\\1&0&0&-1&-1&0&0&0\\0&0&0&-1&1&0&0&1\\0&-1&-1&0&0&0&-1&0\end{bmatrix},B^R=\begin{bmatrix}0&-1&0&0\\-1&0&0&0\\1&0&0&0\\0&-1&0&0\\0&0&0&-1\\0&0&-1&0\\0&0&1&0\\0&0&0&-1\end{bmatrix},$$

and $\operatorname{rank}\left(A^R\right)=\operatorname{rank}\left(A^R,B^C\right)=8$. Clearly, the linear equations $A^RY=B^R$ have a unique solution. The unique solution is easily found to be

$$Y=\frac{1}{9}\begin{bmatrix}6&0&-3&0\\0&-3&0&0\\0&6&0&3\\3&0&0&0\\3&0&6&0\\0&0&0&-3\\0&-3&0&6\\0&0&3&0\end{bmatrix}.$$

By (4.8), we easily find the unique solution X of quaternionic linear equations $AX=B$, and

$$X=\frac{1}{16}\left[I_2,I_2\mathrm{i},I_2\mathrm{j},I_2\mathrm{k}\right]\left(Y-P_nYP_p-R_nYR_p-S_nYS_p\right)\begin{bmatrix}1\\-\mathrm{i}\\-\mathrm{j}\\-\mathrm{k}\end{bmatrix}$$

$$=\begin{bmatrix}\frac{2}{3}+\frac{1}{3}\mathrm{j}&\frac{1}{3}\mathrm{i}\end{bmatrix}^{\mathrm{T}}.$$

Case 2: For split quaternionic linear equations $AX=B$ with $v=1$. It is easy to know A^R and B^R by (2.6b),

$$A^R=\begin{bmatrix}0&1&-1&0&0&1&0&0\\-1&0&0&0&1&0&0&-1\\1&0&0&1&0&0&0&-1\\0&0&-1&0&0&-1&-1&0\\0&1&0&0&0&1&-1&0\\1&0&0&-1&-1&0&0&0\\0&0&0&-1&1&0&0&1\\0&-1&-1&0&0&0&-1&0\end{bmatrix},B^R=\begin{bmatrix}0&-1&0&0\\-1&0&0&0\\1&0&0&0\\0&-1&0&0\\0&0&0&-1\\0&0&-1&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}$$

and $\operatorname{rank}\left(A^R\right)=\operatorname{rank}\left(A^R,B^R\right)=8$. Clearly, the linear equations $A^RY=B^R$ have a unique solution. The unique solution is easily found to be

$$Y=-\frac{1}{5}\begin{bmatrix}-4&0&-1&0\\0&3&0&-2\\0&-4&0&1\\-3&0&-2&0\\-1&0&-4&0\\0&-2&0&3\\0&1&0&-4\\-2&0&-3&0\end{bmatrix}.$$

By (4.9), we easily find the unique solution X of split quaternionic linear equations $AX=B$, and

$$X = \frac{1}{16}\left[I_2, I_2\mathrm{i}, I_2\mathrm{j}, I_2\mathrm{k}\right]\left(Y - P_nYP_p + R_nYR_p + S_nYS_p\right)\begin{bmatrix}1\\-\mathrm{i}\\\mathrm{j}\\\mathrm{k}\end{bmatrix}$$

$$= \begin{bmatrix}\frac{4}{5}+\frac{1}{5}\mathrm{j} & \frac{3}{5}\mathrm{i}+\frac{2}{5}\mathrm{k}\end{bmatrix}^{\mathrm{T}}.$$

5. Conclusion

The goal of this paper is to solve the quaternion and split quaternion linear equations in a unified manner. First, we give the definition of the v-quaternion and two new matrix representations of v-quaternion matrix. Then we derive two algebraic methods for solving the linear equations of v-quaternion. It is noteworthy that this paper not only gives algebraic techniques for solving the linear equations over v-quaternion algebras, but also a unification of algebraic techniques for linear equations in quaternionic and split quaternionic theory.

Acknowledgements

We thank the Editor and the referee for their comments. Research of T. Jiang is funded by the National Natural Science Foundation of China (11771188) and Shandong Natural Science Foundation (ZR201709250116). This support is greatly appreciated.

Conflicts of Interest

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

References

[1] Hamilton, W.R. (1843) On a New Species of Imaginary Quantities Connected with a Theory of Quaternions. *Proceedings of the Royal Irish Academy*, **2**, 424-434.

[2] Özdemir, M. and Ergin, A.A. (2006) Rotations with Unit Timelike Quaternions in Minkowski 3-Space. *Journal of Geometry and Physics*, **56**, 322-336. https://doi.org/10.1016/j.geomphys.2005.02.004

[3] Schmidt, J. and Nieman, H. (2001) Using Quaternions for Parametrizing 3-D Rotations in Unconstrained Nonlinear Optimization. *Vision Modeling, and Visualization*, Stuttgart, 21-23 November 2001, 399-406.

[4] Cockle, J. (1849) On Systems of Algebra Involving More than One Imaginary and on Equations of the Fifth Degree. *Philosophical Magazine*, **35**, 434-437. https://doi.org/10.1080/14786444908646384

[5] Kula, L. and Yayli, Y. (2007) Split Quaternions and Rotations in Semi Euclidean Space E_2^4. *Journal of the Korean Mathematical Society*, **44**, 1313-1327. https://doi.org/10.4134/JKMS.2007.44.6.1313

[6] Özkaldi, S. and Gündoğan, H. (2011) Dual Split Quaternions and Screw Motion in 3-Dimensional Lorentzian Space. *Advances in Applied Clifford Algebras*, **21**, 193-202. https://doi.org/10.1007/s00006-010-0236-6

[7] Özdemir, M., Erdoğdu, M. and Simsek, H. (2014) On the Eigenvalues and Eigenvectors of a Lorentzian Rotation Matrix by Using Split Quaternions. *Advances in Applied Clifford Algebras*, **24**, 179-192. https://doi.org/10.1007/s00006-013-0424-2

[8] Adler, S.L. (1995) Quaternionic Quantum Mechanics and Quantum Fields. Oxford University Press, New York.

[9] Özdemir, M. and Ergin, A.A. (2005) Some Geometric Applications of Timelike Quaternions. *International Conference of the Jangjeon Mathematical Society*, Vol. 16, 108-115.

[10] Ramis, C. and Yayli, Y. (2013) Dual Split Quaternions and Chasles' Theorem in 3-Dimensional Minkowski Space E_1^3. *Advances in Applied Clifford Algebras*, **23**, 951-964. https://doi.org/10.1007/s00006-013-0405-5

[11] Aslan, S. and Yayli, Y. (2016) Split Quaternions and Canal Surfaces in Minkowski 3-Space. *International Journal of Geometry*, **5**, 51-61.

[12] Brody, D.C. and Graefe, E.M. (2011) On Complexified Mechanics and Coquaternions. *Journal of Physics A*: *Mathematical and Theoretical*, **44**, Article ID: 072001. https://doi.org/10.1088/1751-8113/44/7/072001

[13] Janovská, D. and Opfer, G. (2013) Linear Equations and the Kronecker Product in Coquaternions. *Mitteilungen der Mathematische Gesellschaft in Hamburg*, **33**, 181-196.

[14] Jiang, T. (2004) An Algorithm for Quaternionic Linear Equations in Quaternionic Quantum Theory. *Journal of Mathematical Physics*, **45**, 4218-4222. https://doi.org/10.1063/1.1794368

[15] Zhang, Z., Jiang, Z. and Jiang, T. (2015) Algebraic Methods for Least Squares Problem in Split Quaternionic Mechanics. *Applied Mathematics and Computation*, **269**, 618-625. https://doi.org/10.1016/j.amc.2015.07.072

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1732-1754
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

A New Celestial Mechanics Dynamics of Accelerated Systems

Gabriel Barceló

Dinamica Fundación, Madrid, Spain
Email: gestor@advanceddynamics.net

How to cite this paper: Barceló, G. (2019) A New Celestial Mechanics Dynamics of Accelerated Systems. *Journal of Applied Mathematics and Physics*, **7**, 1732-1754.
https://doi.org/10.4236/jamp.2019.78119

Received: July 2, 2019
Accepted: August 13, 2019
Published: August 16, 2019

Abstract

We present in this text the research carried out on the dynamic behavior of non-inertial systems, proposing new keys to better understand the mechanics of the universe. Applying the field theory to the dynamic magnitudes circumscribed to a body, our research has achieved a new conception of the coupling of these magnitudes, to better understand the behavior of solid rigid bodies, when subjected to multiple simultaneous, non-coaxial rotations. The results of the research are consistent with Einstein's theories on rotation; however, we propose a different mechanics and complementary to classical mechanics, specifically for systems accelerated by rotations. These new concepts define the Theory of Dynamic Interactions (TDI), a new dynamic model for non-inertial systems with axial symmetry, which is based on the principles of conservation of measurable quantities: the notion of quantity, total mass and total energy. This theory deduces a general equation of motion for bodies endowed with angular momentum, when they are subjected to successive non-coaxial torques.

Keywords

Mechanics, Dynamic Interactions, Non-Inertial Systems, Dynamic Coupling, Celestial Mechanics

1. Introduction

The observation of the dynamic equilibrium of the universe, together with the verification of the simultaneous orbitation and rotation of the celestial bodies, generated doubts about the accepted rotational mechanics.

This secular dynamic equilibrium did not seem to comply with a Newtonian physics in which the forces generate movements of constantly accelerated translation. The balance that we can observe in our universe, and its dynamics, did

not seem to reconcile with the conceptual structure of Classical Mechanics.

With these doubts we began, from the outset of our speculations about the incoherence of the dynamic orthodox paradigm. First, we conducted a historical study on the concept of rotation in physics, published in the book: *The flight of the Boomerang* [1], whose prologue was written by Professor Federico Garcia-Moliner [2]. Later we published in a new book our initial concerns and hypothesis: *A world in rotation* [3].

After observing the behavior of celestial bodies, we concluded that the application of Newtonian mechanics to mobile systems in non-inertial frames, gave conceptually wrong results.

After analyzing the physical phenomenon of a rigid solid body, free in space, subjected to multiple simultaneous non-coaxial rotations, we concluded that the current dynamics did not adequately justify the behaviors that could be observed.

In making a detailed analysis, it could not be considered that the resulting motion was chaotic; however, the true response of nature was complex and far removed from the accepted argument in the field of Classical Mechanics.

It is clear that this theory determines an ideal model for inertial systems, but is not able to justify moving systems subjected to accelerations, such as all the movements with rotation.

Therefore, it was about analyzing the existing physical-mathematical model for inertial systems and to determine a new modeling for non-inertial systems, establishing its true equation of motion [4].

The challenge was to define a new celestial mechanics based on a Dynamic of Non-Inertial Systems.

2. Incoherences: Identification of Errors

It is hard to believe that to date the true behavior of bodies subject to multiple, non-coaxial rotations has not been identified and that these have been assigned a mathematically unsolvable equation, even in analytical mechanics, which makes it further hard to believe that nature can behave in such a manner: without a resolvable, exact and predictive result. After our analyses, we pointed to the inconsistency in the accepted Newton-Euler equations in their application to different bodies subject to diverse non-coaxial rotations [5].

But it is also astonishing to see that the accepted calculation procedure determines a trajectory for these bodies, which does not coincide at all with what can be observed in nature (see **Figure 1**). This lack of coherence between the results of the formulas applied, and the true trajectory that is observed, makes us think that in all these years, no experimental tests or verifications have been carried out to confirm if those algorithms responded to reality, or if they were simple mathematical structures, alien to the authenticity of nature.

3. Non-Inertial Systems

Classical Mechanics has been formulated for inertial references and it does not

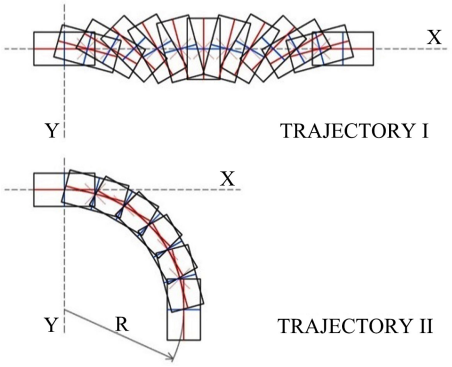


Figure 1. Trajectory I predicted by classical mechanics and trajectory II deduced by means of the Theory of Dynamic Interactions (TDI). We have observed and confirmed the second case repeatedly by experiments [6].

say anything and it does not relate to rotation cases. Nevertheless, creating a mechanics available for any stage is possible if we include their inertial reactions and we define a Fields Inertial Dynamic. In this way the inertial phenomena would be structured rationally, and incorporate a unified mechanics.

In order to include the inertial phenomena in the physics knowledge structure, it is necessary to analyze the movement in non holonomous coordinates and the axial reactions which are produced. We understand that a classical mechanics based on holonomous coordinates and polar reactions, will only represent a limited and partial natural view.

The proposed generalization does not say that the classical mechanics is outdated or wrong, but it that it is limited and partial, because it is only referred to inertial systems. We can be more aspiring and search more general dynamical laws which regulate the movement bodies behavior under rotations or, supporting multiple and non-coaxial rotations.

The Theory of Dynamic Interactions generalizes the gyroscopic couple concept and other inertial phenomena concepts and it includes into a unified structure of the new rotational non inertial dynamic [7].

3.1. Matching Orbiting and Rotation Movements

However, the fact is that, in general, celestial bodies in addition to orbiting also rotate about their main axis of symmetry. Nevertheless, this peculiar feature fails to have attracted the attention of astronomers until our day.

At the beginning of our project, we came to the conclusion that there could be a nomological, physical-mathematical correlation between the simultaneous rotation and orbiting movements that we observe in celestial bodies [4].

In our studies it was confirmed that it is easy to see simultaneous movements of intrinsic rotation and orbitation in nature, when until now there was no physical or mathematical model that established a scientific correlation between both movements [8].

Julio Cano has expressed: By observing in nature the constancy of the relationship

between orbiting and intrinsic rotation, Gabriel Barceló deduced the principle that: Everything that orbits, rotates [9]; or rather, everybody that moves through an orbit simultaneously rotates on an intrinsic axis. He deduced this principle from observing the planetary system, the rings of Saturn and also the behaviour [10] of the spinning top.

He understood, however, the need for empirical checks to confirm or rectify the new dynamic hypotheses deduced from the aforementioned principle and, where appropriate, to be able to explain that behaviour by formulating a new dynamic theory that would simultaneously resolve other Rotational Dynamics phenomena and generalize inertial phenomena [6].

The Theory of Dynamic Interactions allows to justify this constant coincidence between orbiting and intrinsic rotation, and to develop a specific dynamics for bodies in rotation, subjected to successive non-coaxial torques, in which the sequence of the action of the forces is decisive, and their behavior, does not exactly coincide with the laws of Classical Mechanics.

3.2. New Hypotheses

Guided along by our curiosity and the intellectual concerns that our teacher Miguel A. Catalán shared with us, we have tried, over this last sixty years to seek for explanations to our doubts and aporias, trying to transform our conjectural baggage into a logical structure of knowledge to better understand the behavior of nature, to be able to accurately explain it and to come up with the mathematical tools that will make it possible to predict it [5].

To achieve this, it was necessary to observe the singular behavior of the material when subjected to successive accelerations by rotation, not coaxial, and infer their true behavior in non-inertial conditions.

From determinate dynamics assumptions and based on a new interpretation of the bodies with intrinsic angular momentum, when they are supporting successive non-coaxial torques, we have developed new dynamics hypothesis which allows us to reach the conclusion to create a new mathematical model in the rotation dynamics fields.

This new model would allow us to justify dynamic natural behaviors that until now have not been sufficiently understood. With this new conceptual model, new results for determinate cases are obtained based on a new interpretation of the composition of movements originated by external momentums [7].

Our research project was born of physical observation, the search for a dynamic for accelerated systems based on the Scientific Method, and the reflection on the validity of classical mathematical models, which accept to apply vector algebra to angular magnitudes. Therefore, from the beginning, our new differentiating hypotheses were:

- The mathematical model for the development of the classical movement equations admits the vectorial algebra, although the rotational dynamical magnitudes are neither commutative nor associative.

- In Classical Mechanics it is accepted from Poinsot, the discriminant principle of non-overlap between the movements of translation and rotation, and that the effects of the forces do not match the effects of the torques, without this supposed rule had been experimentally tested.

3.3. New Deductions

Dissatisfied as we were with the criteria of conventional rotational dynamics, we tried to come up with new approaches to determine the equation of motion of accelerated systems and to better understand and explain their behavior. After a first analysis, we became convinced that:

- Whenever a solid rigid body with intrinsic rotation is subject to an external, non-coaxial momentum, dynamic interactions are generated that cause a variation in the dynamic behavior of the body, which is not provided for in classical mechanics.
- Solid bodies with intrinsic angular momentum and a quantity of movement, when they are subject to a new, non-coaxial momentum with its rotation, change their path in accordance with a defined regularity [5].

3.4. The New Equation of Motion

Based on the axioms put forward in this deductive process, we came to a new formulation [11] for the movement of accelerated systems: the motion equation would be determined by the translational velocity of the body's center of mass, which has not varied in magnitude and therefore will be equal to the initial translational velocity of the body subjected to the spatial rotation mentioned above:

$$\vec{v} = \vec{\vec{\Psi}} \cdot \vec{V}_0 \tag{1}$$

The no discriminating coupling proposed in my hypothesis is hence identified as a spatial rotation of velocity:

$$\vec{v} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{V}_0 \tag{2}$$

Thus, the equation of motion can be written as

$$\vec{v} = \vec{\vec{\Psi}} \cdot \vec{V}_0 = \begin{pmatrix} \cos\left(M'\right)t/\left(I\omega\right) & -\sin\left(M'\right)t/\left(I\omega\right) & 0 \\ \sin\left(M'\right)t/\left(I\omega\right) & \cos\left(M'\right)t/\left(I\omega\right) & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{V}_0 \tag{3}$$

In a single rotation, the rotational operator $\vec{\vec{\Psi}}$ transforms the initial velocity $\vec{V}_0$ into the velocity $\vec{v}$, both situated in the same plane. We find that the rotational operator $\vec{\vec{\Psi}}$ is a function of sine or cosine of Ωt, which clearly indicates the relation between the angular velocity $\vec{\Omega}$ of the orbit and the torque $\vec{M}'$ and the initial angular velocity $\vec{\omega}$. Thus, I have derived a simple mathematical relation between the angular velocity $\vec{\omega}$ of the body and its translation-

al velocity $\vec{v}$. Equation (3) is a general equation of motion for bodies with angular momentum that are subjected to successive noncoaxial torques. For this equation, the rotational operator $\vec{\vec{\Psi}}$ serves as a matrix that transforms the initial velocity, by means of rotation, into the velocity that corresponds to each successive dynamic state [12].

This new equation of motion makes it possible to learn of the true and real behavior of nature for non-inertial systems. It can be understood as a scientific realization that explains and predicts the behavior of accelerated dynamic systems.

… This new conception of dynamics requires an interpretative sensitivity to rotational phenomena, to which the criteria of translational dynamics cannot be applied. Perhaps it corresponds to that necessary subtle change of perspective, something that we all have misse … [13], referred to by Penrose [5].

3.5. Dynamic Coupling

To arrive at the different results from the physical-mathematical model obtained, we worked off the basis of a new interpretation the concept of dynamic coupling. We put forward new criteria with respect to the compounding or superposition of the movements originated by the acting forces [4].

On the basis of the Principle of Conservation of Momentum, we can infer that the field of inertial reactions generated in the rotating space by a new non-coaxial momentum, upon a moving body with a rotational movement ω and an inertial momentum I upon that rotation axis, and thus with an angular momentum, will oblige the moving body to acquire a precession rate Ω.

This precession rate Ω can be observed simultaneously with the initial ω, which stays constant within the body. Beyond this, and as a discriminant hypothesis, in the case of the translational movement of the body, we propose the dynamic hypothesis of the coupling of this field of translation velocities with the anisotropic field of inertial velocities caused by the second non-coaxial momentum, obtaining as a resultant movement, an orbiting that is simultaneous with the intrinsic rotation of the moving body. This new orbiting movement, generated by a non-coaxial momentum, defines itself through the rotation of the speed vector, the latter being kept constant in module.

Consequently, dynamic effects can be associated with speed, thus highlighting the clear mathematical correlation between rotation and translation. This mathematical correlation allows us to identify a physical relationship between the transfer of kinetic rotational energy to kinetic translational energy, and vice versa [4].

This is the concept of dynamic coupling used in our analysis: they are our working hypotheses. We do not understand that other possible coupling concepts should be analyzed. In this new non-inertial rotational dynamics based on the Theory of Dynamic Interactions, we have developed laws and corollaries [14] (see Treaty), which allow us an unknown number of new technological applica-

tions.

3.6. Brief Summary of the Theory

In the Treaty in two volumes: New paradigm in Physics [15], we propose this new physico-mathematical model for bodies subjected to multiple actions that generate non-coaxial rotations, proposing a new structure of this knowledge.

We also state the laws that determine the behavior of systems in accelerated motion, for example, when they are subjected to non-coaxial moments, and multiple examples of assumptions and phenomena of nature that respond to these laws are proposed.

The Theory of Dynamic Interactions (TDI) explains the behavior of rotating bodies; the reason why, when these bodies are subject to new stimuli that do not share their axis of rotation, an intriguing, different and simultaneous movement is generated. This refers to a behavior that remained unexplained by the conventional laws of physics known until then [16].

Figure 2 illustrates an imaginary field of vectors that represent a supposed dynamic magnitude, for example, the speed of translation of a body, in each of its points.

It is necessary to emphasize that the fields of velocities that are the object of our analysis will be within the mobile under study, and represent that specific magnitude, in each point of the body. If the mobile is subject to a simple translation, without any rotation, the field will be uniform and isotropic, and all the vectors that represent the magnitude of that field will be equal and parallel.

On the other hand, a torque or external moment creates a new velocity field within the mobile. This field will be anisotropic, if the body already has another non-coaxial rotation, since the speed of each point will depend on its geometric position with respect to the axis, in the body that rotates.

Our dynamic hypothesis is that the velocity field of the initial rotation remains constant, but instead, the field of translation speeds is coupled with the field of translation speeds generated by the second rotation, not coaxial.

In this way, after this superposition of fields, the body changes trajectory, and if the external action remains constant, it will describe an orbit.

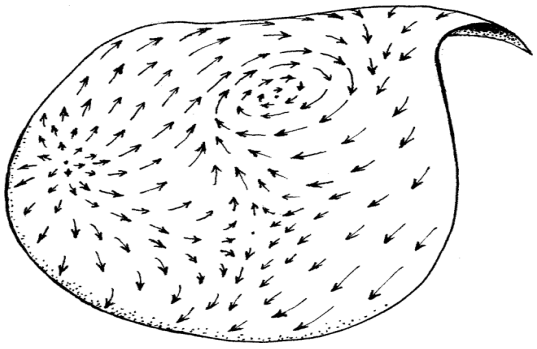


Figure 2. Example of a geometric interpretation of a field of vector dynamic magnitudes generated at the core of the mobile [4].

The conclusion of our proposed theory is that for this assumption, the mobile, will gradually change its trajectory, as shown in **Figure 3** [17].

As we can see in **Figure 3**, we can assume a body with rectilinear trajectory, equipped with translational velocity and intrinsic rotation on its principal axis of inertia, which is subjected to a new non-coaxial moment with the intrinsic rotation, for example in the figure, a flotation/weight torque, contained in the drawing plane, as in our submarine experiment [18] [19] [20].

In such case, the anisotropic velocity field, generated by this flotation/weight torque, forces the mobile to rotate about a vertical axis, perpendicular to that of the external torque acting on it (See **Figure 3**). In red is the displaced mobile, but with the previous orientation, and in blue, the new orientation of the mobile, due to the dynamic coupling taking place. The result is the coupling of both the translational and the anisotropic fields and, consequently, the change in the mobile trajectory, describing an orbit, if its initial velocity was constant [4].

4. Experimental Tests

This new logical structure was contrasted with experimental tests and computer simulation models, obtaining a full coherence between the results of the simulations and the observation of the results of the empirical evidences. These tests were carried out by the Advanced Dynamics research team, but also by independent third parties [21], who designed their own experimental testing prototypes [22].

These tests can be visualized in the videos that appear in the annex I. The last video made [23], shows one of the multiple experimental examples of the Theory of Dynamic Interactions, but it is possibly one of the simplest and most striking: a cylinder operated by a finger, which acts on one of its edges. When you push the boat or cylinder with your finger, you rotate it on its longitudinal axis (main torque), and simultaneously also, on a vertical axis (secondary torque).

In the video you can see the velocity fields that are generated. There is a main angular momentum (shown with arrows), a secondary angular torque perpendicular to the previous one (shown with arrows), and a rectilinear velocity of the

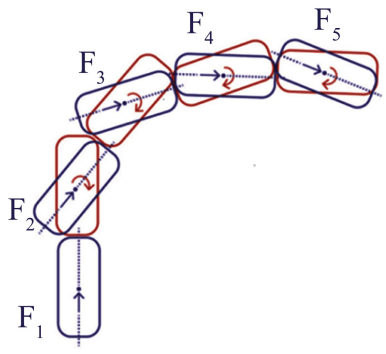


Figure 3. Trajectory of a body equipped with translational speed and intrinsic rotation on its principal axis of inertia, when it is subjected to a new non-coaxial moment with the intrinsic rotation.

center of mass of the boat (shown with arrows). Both rotations, and the translation of the center of mass of the boat cause it to rise, without the need for an external force.

The boat rises with a tendency to become vertical; it can even stand supported on its base, in a stable position on the flat surface of the ground, but without the need for any external force acting in that direction. When the second rotation enters into action, the velocity distribution in the cylinder particles no longer remains constant, but is variable, in accordance with what is stated in the Theory of Dynamic Interactions. The generation of a variable velocity field implies the appearance of accelerations.

The velocity field generated by the second pair will have a vertical component that is coupled with the field of the speed of translation, and forces the center of mass of the cylinder to rise, that is, the movement observed is opposite to the action of the weight on the cylinder, whose tendency would be to fall. This video is an evident proof of the theory that is sustained, and we recommend its visualization: https://www.youtube.com/watch?v=hJSbVOHRfrU.

These experimental references, and many others that may arise, infer the existence of another rotational dynamics, not Newtonian, necessary for the identification of the behavior of bodies in rotation, when they are subjected to new non-coaxial stimuli, and to which their behavior. Currently, in many cases, it is understood as anomalous, paradoxical or chaotic, since the laws available to us do not allow us to identify and predetermine it [24].

4.1. Verification and Falsifiability

We understand that, whatever the case, our theory, as an alternative to the other classical paradigm, must be subject to independent checks and tests, owing to the clear conceptual discrepancy that exists between them. We believe it necessary to revise this niche of knowledge from a new mindset; one which discards preconceived ideas of a translational scenario and accepts the rotational reality of matter and our environment. Indeed, that has been the main reason that led us to establishing more general axioms to analyze the dynamics of bodies subject to accelerations and, especially, to acceleration by rotation, and to propose a mechanics for non-inertial systems [5].

Both the experimental tests carried out are easily reproducible according to the scientific method. Advanced Dynamics has convened three successive competitions for the possible refutation or antithesis of the proposed theory, without obtaining an answer. The last call for antithesis ended on June 15, 2019 [25].

4.2. Justification

Through repeated experimental tests [26], they reliably saw how certain accepted mathematical formulations do not faithfully show the true dynamic behavior of bodies subject to accelerations by simultaneous non-coaxial rotations [8].

On the contrary, the sustained theory justifies and explains many scientific concepts that could not be understood using Classical Mechanics.

TDI fully explains the planar orbits Kepler proposed. In the laws of classical mechanics, there is no mathematical correlation which relates the movements of orbit and rotation of the planets around the Sun. However, the question of the existence of a physical relation between both movements, which has to date not been mathematically shown can be raised; … the TDI can even explain it.

For example, TDI also explains how flat rings with multiple satellites are often formed in our solar system, such as Saturn ones … Newton was also unable to explain the reason why Saturn's rings, and the many other ring systems in our Solar System, are flat.

Our theory can also explain the reason for its formation and even the ripples that occur in Saturn rings, and all the ring systems in general, owing to changes or disturbances in the acting momenta [4].

Also, the same reasoning could be applied to understand the behavior of many solid elements in rotation such as the boomerang [27] [28], spinning top [29], hoop, gyroscopic pendulum [30], or wheel. The Advanced Dynamics team has made multiple analyzes of these behaviors, especially in the case of the boomerang [31], having published publishing some of the results of these investigations [32].

5. New Paradigm

In the Treaty New paradigm in Physics, as a summary of our investigations, a new mechanical theory has been proposed for bodies subjected to accelerations; also exposing its content, and remembering how this new niche of scientific knowledge has been managed.

After this process of deduction and inference, we have built a knowledge structure for non-inertial systems that incorporates a causal demonstration of accelerated phenomena, without having knowledge to date, as we have stated, of any explicit refutation or antithesis to our reasoning. By means of a pondered management of this knowledge, based on established principles and axioms, we have structured our end thesis, which we call the Theory of Dynamic Interactions (TDI).

We believe that our theory provides a clear, satisfactory and scientific explanation of the rotational phenomena of bodies with axial symmetry [5].

Numerous articles [33] and papers have been published in congresses [34]. Also a Web on the content of the treaty: https://newparadigminphysics.com.

Different videos have been developed to describe its content:
https://www.youtube.com/watch?v=GapMJEfHJjUç;
https://www.youtube.com/watch?v=45kUpdAXICw;
https://www.youtube.com/watch?v=Gbx5wdQqTTs;
https://www.youtube.com/watch?v=Gbx5wdQqTTs.

6. Innovations and Conclusions

The result of our research project is the proposal of multiple conceptual and technological innovations for accelerated systems, based on the new equations,

which we have deduced for the movement accelerated by rotations, in non-inertial dynamics.

We maintain that, in the case of systems subject to accelerations, the laws of Classical Mechanics, foreseen for inertial systems, cannot be applied, and new axioms and laws of dynamic behavior, which have been enunciated, are necessary.

6.1. Renovating Proposals

Below we list some of the ideas expressed in this work, that we believe could prove innovative:

1) Non-discriminating coupling

As a result of the new axioms and proposed laws, we deduce that in rotational dynamics, it is necessary to admit the non-discriminative coupling of dynamic magnitudes expressed in Section 3.4, which is represented in Figure 1, and which generates the orbital movement, simultaneous with the intrinsic rotation of the mobile.

But this dynamic, we understand that it is applicable to the cosmos, since in it, celestial bodies, in addition to having movement of translation, rotate and orbit. In addition, the rotational inertia will maintain that constant rotation, while other external actions do not prevent it: This is the reason for the secular rotation of the celestial bodies. Any new non-coaxial external action will generate orbiting movements, and not a translational expansion.

2) Law of simultaneity of orbiting and rotation movements

In our investigation we have observed and enunciated a regularity in the behavior of the celestial bodies, proposing a law of simultaneity of orbit and rotation. We reiterate that the Theory of Dynamic Interactions allows to justify that constant coincidence between orbiting and intrinsic rotation.

3) Immutable dynamic balance

One of the characteristics of our new conception of the cosmos is its constant and lasting dynamic equilibrium, because of the real behavior of matter, when bodies are endowed with intrinsic rotation.

It is the balance that the human being has perceived, when observing the celestial dome, for millennia.

4) Flat celestial systems and Spiral structures

The theory justifies the reason for the formation of rings, and in general of all the celestial plane systems that we observe in our universe.

If we look closely at the universe, we see that it is constituted by flat celestial systems. The galaxies have a flat structure, solar systems like ours, Saturn's rings, the asteroid belt, Kuiper's belt or the Scattered Disk. Precisely, the dynamic model of the TDI, justifies a flat image of the universe, in which galaxies and systems also tend to develop in flat structures [36].

The theory also justifies spiral structures, such as the arms of galaxies (See Figure 4), or tornadoes [37]. It can be understood that these structures are the expression of numerous bodies with accelerated speed and intrinsic rotation,

subjected to non-coaxial moments.

5) Oriented point

In Classical Mechanics the abstract concept of "point" of Euclidean geometry is used, however, this concept is not ideal, and it is not enough to identify the behavior of bodies in rotational dynamics.

It is necessary to accept the use of a new concept of oriented point, to be able to integrate the inertial behavior of the matter in rotational dynamics, and to incorporate, in addition to its situation, and its orientation, the variation of this. This oriented point of the rotational dynamics is identified with 10 coordinates: x_1, x_2, x_3, φ_1, φ_2, φ_3, θ_1, θ_2, θ_3, and time t, thus having 10 degrees of freedom.

6) Fictitious forces and phenomenological effects

In our research project we have repeatedly rejected the conceptual use of *fictitious forces*. We do not understand that in the structure of mechanical knowledge fictitious, apparent or supposed forces must be incorporated, nor that the observed behaviors can be justified with phenomenological examples, such as the gyroscopic effect.

To the current Classical Mechanics, it has been adding unstructured and phenomenological effects, which are not deduced from the initial axioms. We understand that Mechanics must constitute a structure of complete knowledge, without it being necessary to explain it with phenomenological effects, or with fictitious forces.

TDI justifies gyroscopic phenomena by field theory. The gyroscopic torque and the gyroscopic reactions are not a special and singular phenomenology of mechanics, but part of the deductive structured knowledge, of the non-inertial mechanics of accelerated systems.

7) Adding angular momenta

The increase in angular momentum may generate a new movement, different and simultaneous with the existing one, which we will call the precession movement, see Figure 5. We can remember the words expressed by Alejandro Álvarez Martínez: On observing the behavior of moving objects in rotation, when they are exposed to new non-coaxial rotations, Dr. Barceló reaches the conclusion that angular movements of non-coaxial momenta do not necessarily couple [36].

This would be the case of the top and the gyroscope, and in the case of a body in space, without ligatures such as the boomerang, the body will initiate an orbit, without the necessary existence of a central force.

8) Physical-mathematical model: Inertial reactions

In the rotational dynamics, inertial reactions within the mobile are generated, defined and regular, which have been incorporated into the logical and scientific structure of the TDI.

These reactions are incorporated by analyzing the fields of the generated dynamic magnitudes. For example, the field of translation speeds that each external action generates in the body is determined, and the real coupling of the fields that occur in nature is allowed.

Figure 4. Spiral galaxy NGC 6814, whose arms, according to the provisions of the TDI, will be constituted by celestial bodies in rotation and accelerated translation [35].

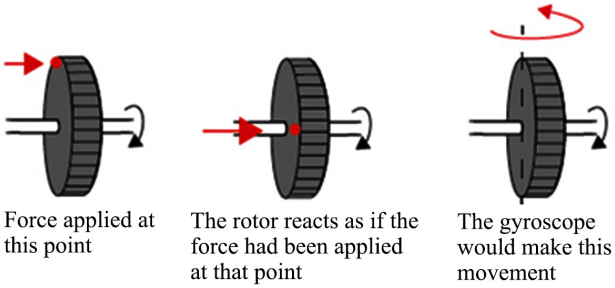


Figure 5. Precession is the response of the mobile to any external action non-coaxial with its own spin. The consequence of this reaction is that the mobile with previous intrinsic rotation modifies the result of the action, apparently displacing the point of application of the force by 90˚, in direction of rotation of the object [17].

From these concepts, a new physical-mathematical model in rotational dynamics has been designed to understand the behavior of nature under accelerated conditions, including simultaneous non-coaxial rotations. This model is complemented with axioms, laws and corollaries, and with a new deduced movement equation.

The initial dynamic hypotheses, and the new physico-mathematical model proposed in rotational dynamics, were confirmed with the experimental tests carried out.

9) Energy transfer

The TDI confirms that it is possible to associate dynamic effects to speed and justifies a clear mathematical correlation between rotation and translation. This mathematical coupling allows us to identify a physical relationship between the kinetic rotational energy transfers to kinetic energy of translation, and vice versa.

10) Contributions to the Theory of Relativity

According to the General Theory of Relativity, we can estimate that the mass of the Earth distorts space-time in its surroundings. In this case, we can assume the analogy that the Moon makes a rolling movement on the curved surface of

the space-time deformed by the Earth, generating a new rotation of the satellite, which we can suppose is not coaxial to the intrinsic rotation that it already has.

In this case, the dynamic interactions predicted by the TDI would be generated, resulting in the closed and flat orbit of the Moon that we see. In this way, we justify the behavior of the celestial bodies, in accordance with the criteria of relativity, without needing torques or forces.

In this same area, the second Law of Kepler can also be justified, since, in the case of an elliptical orbit, it must have a cause according to the TID, in a variation of the orbital velocity, which is consistent with the greater distortion of space-time in the vicinity of the central mass [38].

However, according to the TDI: We suggest that an observer can identify whether or not a body was previously in a state of absolute rest or absolute absence of rotation, and this leads us to the conclusion that movement does not necessarily have to be a relative concept.

This all leads us to propose that the Equivalence Principle is fully valid for the situation put forward by Albert Einstein, but cannot be generalised to all other cases of moving bodies [39].

We propose a revision of the mathematical models deduced from the General Theory of Relativity, incorporating the dynamic criteria of the TDI and its equation of the movement.

11) Planck constant

The concept of Planck radiation quantization, and its constant, understood as the smallest amount of energy that can be transmitted, could be related to the notion of spin, and be explained with the help of the TDI.

We have even proposed that this minimum value of the physical quantities could be justified by the value of the angular momentum of the atomic particles. We suggest a detailed analysis of this proposal, even applying a geometric analogy, and assuming an intrinsic rotation speed of these atomic particles, in such a way that the constant is proportional to the time necessary for a complete rotation, or even a fraction of rotation.

Therefore, TDI … can influence the quantization concept of Planck radiation, and its constant, … We even proposed that this minimum value of physical quantities can be justified by the angular momentum value of atomic particles [17]. This could be a new line of research based on the results obtained in this project.

12) A new Celestial Mechanics

It is thus, according to the TDI, an observable universe in which we find ourselves, with celestial bodies in constant rotation and describing orbits in space, but maintaining a dynamic equilibrium, except for exceptional situations.

This new vision of the universe, we exposed it in Imago Universi [40], and in the article: Proposal of New Criteria for Celestial Mechanics [41]. It is a new conception of Celestial Mechanics based on dynamic non-inertial hypotheses, for bodies accelerated by rotations.

6.2. Applications

In previous texts, we have proposed that, through this analysis, the nature of any movement in space can be determined and predicted, defining its relativity. The movement equation that is proposed, and the laws that are formulated, permit the initiation of the structuring of a rational mechanics and of a rotational dynamics based on principles and axioms, for bodies submitted to accelerations by rotations, clearly differentiated from classical mechanics.

In this new rational structure, phenomena that are paradoxical or alien to the main structure should not be present as happens in classical mechanics with the so called gyroscopic torque or fictitious forces.

The Theory of Dynamic Interactions is a logical-deductive system constituted from some dynamic hypotheses. By means of the observation of nature, the establishing of some initial hypotheses, and starting from axioms and postulates, we have constructed a structure of knowledge in relation to rigid solid bodies, when submitted to successive accelerations by rotation.

The physical-mathematical model obtained allows us to interpret the observable behavior of these bodies, subject to successive non coaxial torques, according to deduced laws, as well as to extract new consequences, inferences and predictions. For example, the theory allows justifying the deviation that undergoes the horizontal curvilinear trajectory of a ball, when it is submitted to non-coaxial moments (See **Figure 6**) [38].

This text and the videos that accompany it, provide only a brief summary of the work and studies carried out to propose a Rotational Dynamics of Interactions applicable to accelerated bodies, even those subjected to multiple successive moments of force, not coaxial [42].

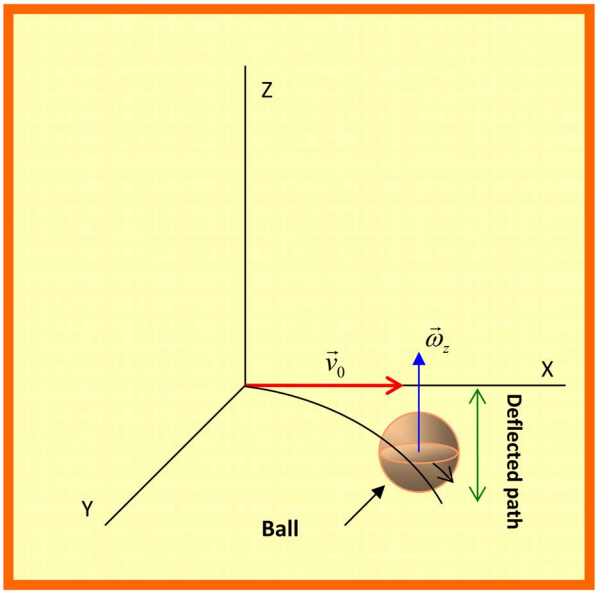


Figure 6. Horizontal curvilinear trajectory of a ball with effect, whose deviation can be justified by the Theory of Dynamic Interactions [4].

6.2.1. Scientific Applications

The **Theory of Dynamic Interactions** can have a transcendental effect on the foundations of dynamics [43], but also on mechanics, atmospheric dynamics, atomic physics, etc. It generalizes the concept of gyroscopic torque, and that of other inertial phenomena, incorporating them into the unified structure of a new non-inertial rotational dynamics.

Through this model of dynamic interactions, one could justify how a body in rotation can initiate an elliptical, circular or even helical trajectory, without the existence of a true central force. According to this dynamic model, the application of a torque of forces to a body with intrinsic rotation, generates a stable system, and in constant dynamic equilibrium.

The result of this project is the conception of an innovative dynamic, but also the demonstration of a rational theory of fields, which allows a new understanding of the behavior of matter [38]. The application of these dynamic hypotheses to astrophysics [44], astronautics [45] and other fields of physics and technology will possibly allow new and stimulating advances in research [46].

6.2.2. Technological Applications

This theory can also have numerous technological applications in the control of movement, in astronautics, in nuclear fusion plants [47] or to interpret climatological phenomena with masses of fluids in rotation, such as typhoons or tornadoes [37].

With the help of the Theory of Dynamic Interactions, a ship without a spade rudder, with energy savings, or a spaceship could be governed [48]. This theory can also be applied in the confinement of plasma, in nuclear fusion reactors [49].

Also has numerous and significant … applications, especially in orbital dynamics, orbit determination, and orbit control. For instance:

- Variation of the affecting torque, arises when subjecting intrinsic angular momentum bodies to new non-coaxial momentums.
- To conceive an intrinsic rotating mobile solid, which could be exclusively controlled due to Dynamic Interactions.
- To calculate the trajectory of any intrinsic angular momentum solid in space.
- To propose a new steering system independent from a rudder or any other external element [38].

We can suggest advances in the studies and application related to orbital mechanics, guidance, navigation, and control of single or multi-spacecraft systems as well as space robotics and rockets [24].

6.3. Conclusions

It is necessary to admit the existence of a rotational dynamics of interactions with real results and which modifies the behavior of bodies in accordance with some specific and universal dynamic Laws.

We want to suggest that interest should arise in physics in the exploration of non-inertial accelerated systems, and also to express a call for the need to devel-

op scientific investigation projects for their evaluation and analysis, as well as technological projects based on these hypotheses. In our opinion, these hypotheses suggest new keys to understand the dynamics of our environment and the harmony of the universe. A universe composed not only of forces, but also of their momentums; and when these act constantly upon rigid rotating bodies, with an also constant translation speed, the result is a closed orbiting movement, thus a system which is moving, but within a dynamic equilibrium.

The application of these dynamic hypotheses to astrophysics, astronautics and to other fields of physics and technology possibly allows new and stimulating advances in investigation.

The result of this project is the conception of an innovative dynamic theory, which specifically applies to rigid rotating physical systems and which has numerous and significant scientific and technological applications.

The Theory of Dynamic Interactions establishes new conceptual criteria, of a more general description, to understand the behavior of nature, meaning that the current laws of dynamics could be considered special and specific cases of this theory. For example, Newton's laws would apply to the case of a physical model of behavior, without force momentums [38].

We have also referred to examples in nature, which support the Theory of Dynamic Interactions, supporting the proposed laws [14].

Next, we will list, a series of characteristics of the physical theory that is proposed. It can be understood that this relationship is a proposal of conclusions, without pretending to be exhaustive:

1) Criteria

The initial hypotheses of the theory are based on new interpretative criteria, for example, on speed coupling and rotational inertia, and on a new mathematical model, which allows the simulation of the real behavior of the bodies subjected to these excitations.

As we have already expressed, our hypotheses have been confirmed by experimental tests.

2) Content

The Theory of Dynamic Interactions is a logical-deductive ordering of knowledge, based on certain dynamic hypotheses. Through the observation of nature, the establishment of initial hypotheses, and the statement of principles, axioms and postulates, we have constructed a structure of knowledge in relation to rigid solid bodies, when subjected to successive accelerations by rotation.

The mathematical physical model obtained allows us to interpret the observable behavior of these bodies, subjected to successive non-coaxial pairs, according to deduced laws, as well as to extract new consequences, inferences and predictions.

3) Harmony of the Universe

We have tried to imagine the trajectory of a body in rotation in space, when it is forced to make new rotations on new axes, obtaining as a result the orbital iti-

nerary that we observe in the celestial bodies, and a balance and secular harmony of our universe. A universe that is not necessarily expanding, according to the TDI [43].

4) Field theory

The result of this research project is the conception of an innovative dynamic, and the exposition of a rational field theory, which provides a new understanding of the behavior of matter.

Our research can be expanded, in addition to the study of Field Theory, with a deep relativistic analysis, which allows knowledge of the dynamics of new physical systems of space, and the potential future achievement of numerous and relevant new technological developments.

5) Consistency with classical mechanics

This theory does not intend to challenge Newton's laws; since with them a conceptual structure of proven reliability has been developed. We propose a different and complementary mechanics to Classical Mechanics, specifically for systems accelerated by rotations. Therefore, both theories do not come into conflict, as they do not refer to the same domain of knowledge. They are theories of specialization, in their respective fields of action.

We propose a new Celestial Mechanics based on a specific rotational algebra for non-inertial areas, in which the initial hypotheses of the laws of the translational movement of Classical Mechanics are not respected. We propose the exploration of a new knowledge niche for specific dynamic conditions, but not trivial, because they can be observed repeatedly in our universe.

It is necessary to admit the existence of a rotational dynamics of interactions, with real results, and that modifies the behavior of bodies according to universal dynamic laws that were not known to date.

6) Correlation

In our research project, a total coherence can be found between the initial speculations, the original hypotheses, the principles and axioms applied, the physical laws deduced, even causal laws that justify the observed behavior, the mathematical physical models corresponding to the equations of movement which result from the dynamic laws deduced, the simulation models achieved, and the experimental tests carried out. There are videos with these experimental tests (See **Appendix**).

We believe that with this new model that we propose, it will facilitate the understanding of our observational universe, and of the physical phenomena that we notice in it.

We want to suggest the interest that the exploration of accelerated non-inertial systems should have in physics, and express a call to the need to develop scientific research projects in this field, for their evaluation and analysis, as well as technological projects based on these hypotheses.

To obtain more information about this theory, we suggest going to the books and texts referred to and also visiting the following portals:

https://newparadigminphysics.com/
http://www.advanceddynamics.net/
http://www.dinamicafundacion.com/
http://www.tendencias21.net/fisica/
https://club.tendencias21.net/mundo/
http://imagouniversi.com/

Conflicts of Interest

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

References

[1] Barceló, G. (2006) The Flight of the Boomerang. Ed. Marcombo, Barcelona.

[2] Garcia-Moliner, F. (2017) Physical-Mathematical Models in Rotational Motions. *World Journal of Mechanics*, **7**, 35-38. https://doi.org/10.4236/wjm.2017.73004 http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=74661

[3] Barceló, G. (2008) A Rotating World. Editorial Marcombo, Barcelona. http://www.dinamicafundacion.com

[4] Barceló, G. (2018) New Paradigm in Physics, Volume II: Assumptions and Applications of the Theory of Dynamics Interactions.

[5] Barceló, G. (2017) New Paradigm in Physics, Volume I: Theory of Dynamics Interactions. Agencia del ISBN, Madrid.

[6] Cano, J. (2015) The Pendulum of Dynamic Interactions. *Journal of Applied Mathematics and Physics*, **3**, 1186-1198. http://www.scirp.org/journal/jamp https://doi.org/10.4236/jamp.2015.39146

[7] (2016) Research Blog: Extensive Report on the Investigations of Gabriel Barceló. *Global Journal of Researches in Engineering*. http://blog.gjre.org/2016/03/behaviour-of-rotational-bodies.html

[8] Barceló, V. (2017) A Scientific Legacy: Theory of Dynamics Interactions. *World Journal of Mechanics*, **7**, 85-100. https://doi.org/10.4236/wjm.2017.73009 http://www.scirp.org/Journal/Home.aspx?IssueID=9235#74661

[9] Barceló, G. (2006) The Flight of the Boomerang. Ed. Marcombo, Barcelona, 311. http://www.dinamicafundacion.com

[10] http://www.advanceddynamics.net/spinning-top-video

[11] Barceló, G. and Álvarez, A. (2007) Mathematical Model in Rotational Dynamics. International Congress of Mathematics in Engineering and Architecture. Polytechnic University of Madrid, T. S. Civil Engineering, Madrid. http://advanceddynamics.net/congresos-reuniones-y-conferencias

[12] Barceló, G. (2012) Analysis of Dynamic Fields in Non-Inertial Systems. *World Journal of Mechanics*, **2**, 175-180. https://doi.org/10.4236/wjm.2012.23021

[13] Penrose, R.R. (2006) The Road to Reality: A Complete Guide to the Laws of the Universe. Débate, Barcelona.

[14] Barceló, G. (2013) Theory of Dynamic Interactions: Laws of Motion. *World Journal of Mechanics*, **3**, 328-338. https://doi.org/10.4236/wjm.2013.39036

[15] https://www.amazon.com/dp/846979518X/ref=rdr_ext_tmb

[16] Dalby, F. (2017) Rolling Over into the Age of Algorithm. *World Journal of Me-*

chanics, **7**, 39-42. https://doi.org/10.4236/wjm.2017.73005
http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=74663

[17] Barceló, G. (2018) Theory of Dynamic Interactions. *Global Summit on Physics Congress*, Madrid, 27 September 2018, 1-31.
https://dinamicafundacion.com/wp-content/uploads/2018/10/GSP2018ENG.pdf

[18] https://www.youtube.com/watch?v=dtMqGSU9gV4&feature=related

[19] https://www.youtube.com/watch?v=qK5mW2j2nzU&feature=related

[20] https://www.youtube.com/watch?v=vSUkd4slHGQ

[21] https://vimeo.com/68763196

[22] Pérez, L.A. (2013) New Evidence on Rotational Dynamics. *World Journal of Mechanics*, **3**, 174-177. http://www.scirp.org/journal/wjm
https://doi.org/10.4236/wjm.2013.33016

[23] Pérez, L.A. (2018) Cylinder Subjected to Two Non Coaxic Rotations.
https://www.youtube.com/watch?v=hJSbVOHRfrU

[24] Barceló, G. (2010) On the Equivalence Principle. In: 61*st International Astronautical Congress*, American Institute of Aeronautics and Astronautics, Prague.

[25] http://www.coiim.es/forocientifico/FORO%20CIENTFICO/Documentos/ON_THE_EQUIVALENCE_PRINCIPLE.pdf

[26] https://www.youtube.com/watch?v=P9hGgoL5ZGk&feature=related

[27] Barceló, G. (2014) Theory of Dynamic Interactions: The Flight of the Boomerang. *Journal of Applied Mathematics and Physics*, **2**, 569-580.
https://doi.org/10.4236/jamp.2014.27063
http://www.scirp.org/journal/jamp

[28] Barceló, G. (2015) Theory of Dynamic Interactions: The Flight of the Boomerang II. *Journal of Applied Mathematics and Physics*, **3**, 545-555.
https://doi.org/10.4236/jamp.2015.35067
http://www.scirp.org/journal/jamp

[29] Alejandro, Á.M. and Almudena, M.G. (2016) The Dance of the Spinning Top. *Global Journal of Science Frontier Research*: *A Physics & Space Science*, **16**, 15-27.
https://globaljournals.org/GJSFR_Volume16/E-Journal_GJSFR_(A)_Vol_16_Issue_3.pdf

[30] https://www.dropbox.com/s/rrjb1786ub75a8h/PIDing_m.mp4?dl=0
http://www.advanceddynamics.net/the-pendulum-video

[31] https://vimeo.com/129383447

[32] Almudena, M.G. (2017) The Flight of the Boomerang: Comments. *World Journal of Mechanics*, **7**, 46-67.
http://www.scirp.org/Journal/Home.aspx?IssueID=9235#74661

[33] Merino, J. (2017) The Works and Days of Gabriel Barceló. *World Journal of Mechanics*, **7**, 43-45. https://doi.org/10.4236/wjm.2017.73006
http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=74664

[34] Barceló, G. (2009) Dinámica de sistemas acelerados por rotación XXXII Reunión Bienal de la R.S.E. de Física. Ciudad Real, septiembre.
http://www.advanceddynamics.net

[35] ESA/Hubble & NASA (2017) Acknowledgement: Judy Schmidt.
https://www.nasa.gov/image-feature/goddard/2016/hubble-spies-a-spiral-snowflake

[36] Alejandro, A.M. (2017) Theory of Dynamic Interactions: Innovations. *World Journal of Mechanics*, **7**, 101-119. https://doi.org/10.4236/wjm.2017.73010

[37] Barceló, G. (2014) Dynamic Interactions in the Atmosphere. *Atmospheric and Climate Sciences*, **4**, 828-840. https://doi.org/10.4236/acs.2014.45073 http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=51584#.VHB4YTSG_To

[38] Barceló, G. (2017) Theory of Dynamic Interactions: Synthesis. *Transactions on Machine Learning and Artificial Intelligence*, **5**, 10-25. http://scholarpublishing.org/index.php/TMLAI/article/view/3344

[39] Barceló, G. (2014) On Motion, Its Relativity and the Equivalence Principle. *Journal of Modern Physics*, **5**, 1839-1847. https://doi.org/10.4236/jmp.2014.517180 http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=51422#.VHB0jzSG_To

[40] Barceló, G. (2013) Imago Universi. Arpegio, Barcelona. https://dinamicafundacion.com/imago-universi/?lang=en

[41] Barceló, G. (2013) Proposal of New Criteria for Celestial Mechanics. *International Journal of Astronomy and Astrophysics*, **3**, 385-391. https://doi.org/10.4236/ijaa.2013.34044

[42] Cano, J. (2017) Rotational Dynamics: A Challenge. *World Journal of Mechanics*, **7**, 68-84. http://www.scirp.org/Journal/Home.aspx?IssueID=9235#74661

[43] The Universe Does Not Necessarily Expand Forever. https://www.alphagalileo.org/en-gb/Item-Display/ItemId/161784?returnurl=https://www.alphagalileo.org/en-gb/Item-Display/ItemId/161784

[44] Barceló, G. (2006) A New Rotational Dynamics of Interactions for the Planet Saturn. http://dinamicafundacion.com/wp-content/uploads/2014/02/UNA-NUEVA-DINAMICA-ROTACIONAL-DE-INTERACCIONES-PARA-EL-PLANETA-SATURNO.pdf

[45] Barceló, G. (2005) Dynamic Anomalies in the Pioneer Space Probes. RSEF's 2005 XXX Biennale. http://dinamicafundacion.com/wp-content/uploads/2014/02/ANOMAL%C3%8DAS-DIN%C3%81MICAS-EN-LAS-SONDAS-PIONEER1.pdf

[46] Barceló, G. (2011) Analysis of Dynamics Field Systems Accelerated by Rotation. Dynamics of Non-Inertial Systems. *DeMSET*-2011 *Congress*, Miami, 30 November 2011. http://www.coiim.es/forocientifico/FORO%20CIENTFICO/Documentos/DeMSET_2011_GBarcelo.pdf

[47] Barceló, G. (2014) Dynamic Interaction Confinement. *World Journal of Nuclear Science and Technology*, **4**, 249-260. https://doi.org/10.4236/wjnst.2014.44031 http://www.scirp.org/journal/PaperInformation.aspx?paperID=51026&

[48] Barceló, G. (2013) Technological Applications of the New Theory of Dynamic Interactions. *Global Journal of Researches in Engineering*: *Mechanical and Mechanics Engineering-G*, **13**, 33-39. https://globaljournals.org/GJRE_Volume13/E-Journal_GJRE_(G)_Vol_13_Issue_5.pdf

[49] Barceló, G. (2016) Dynamic Interaction: A New Concept of Confinement. *Global Journal of Science Frontier Research*: *A Physics & Space Science*, **16**, 1-9. https://globaljournals.org/GJSFR_Volume16/E-Journal_GJSFR_(A)_Vol_16_Issue_3.pdf

Appendix

Experimental Tests and Videos

Several test experiments have been conducted over recent years, the results of which have been fully satisfactory. These tests serve to confirm the dynamic hypotheses that underpin the Theory of Dynamic Interactions. Videos have been recorded of these tests that can be seen at the following links:

Barceló, G.: *Theory of Dynamic Interactions*. Videos, 2002.

http://www.youtube.com/watch?v=P9hGgoL5ZGk&list=PL3E50CF6AEBEED47B

http://www.youtube.com/watch?v=XzTrGEtJGXU&list=PL3E50CF6AEBEED47B

http://www.youtube.com/watch?v=dtMqGSU9gV4&list=PL3E50CF6AEBEED47B

http://www.youtube.com/watch?v=qK5mW2j2nzU&list=PL3E50CF6AEBEED47B

Bauluz, E.: *New Dynamic Hypotheses*. Madrid, 2011. This video showed the experiments carried out by Advanced Dynamics S. A. to prove and justify the http://www.youtube.com/watch?v=vSUkd4slHGQ

Sanchez Boyer, J.: *Imago Universi*. Video, Madrid, 2013.

https://vimeo.com/62247544

Pérez, L. A.: Reflecting *New Evidence on Rotational Dynamics*, 2013. Video.

http://vimeo.com/68763196

Sanchez Boyer, J.: *The Flight of the Boomerang* II, Video. 2015.

https://www.youtube.com/watch?v=mGfrGW5fhOg&feature=youtu.be

https://vimeo.com/129383447

Pérez, L. A.: *The Pendulum of Dynamic Interactions*. Video. 2015.

https://vimeo.com/160873005

Pérez, L. A.: *The Dance of the Spinning Top*. Video, Valladolid, 2015.

www.advanceddynamics.net/spinning-top-video/

Pérez, L. A.: *Cylinder Subjected to Two Non Coaxic Rotations*. 2018.

https://www.dropbox.com/s/wgb6oztjvdcnziy/EBP_EN.mp4?dl=0

Web Advanced Dynamics: Animations and Videos

http://advanceddynamics.net/en/medios-audiovisuales/

3.0 Submarine Simulation.

3.1 Submarine Prototype I.

3.2 Submarine Anisotropic Field.

3.3 Resultant Field.

3.4 Prototype II.

3.5 Prototype II Simulation.

3.6 Prototype Prototipo II navigation.

3.7 Catamaran.

4.0 Translation Velocity Field.

4.1 Generating a New Field.

Journal of Applied Mathematics and Physics, 2019, 7, 1755-1774
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Stability of High-Order Staggered-Grid Schemes for 3D Elastic Wave Equation in Heterogeneous Media

Atish Kumar Joardar[1,2,3], Wensheng Zhang[1,2*]

[1]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China
[2]School of Mathematical Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China
[3]Department of Mathematics, Islamic University, Kushtia, Bangladesh
Email: *zws@lsec.cc.ac.cn, atishjoardar@gmail.com

How to cite this paper: Joardar, A.K. and Zhang, W.S. (2019) Stability of High-Order Staggered-Grid Schemes for 3D Elastic Wave Equation in Heterogeneous Media. *Journal of Applied Mathematics and Physics*, **7**, 1755-1774.
https://doi.org/10.4236/jamp.2019.78120

Received: July 4, 2019
Accepted: August 16, 2019
Published: August 19, 2019

Abstract

In this paper, we firstly derive the stability conditions of high-order staggered-grid schemes for the three-dimensional (3D) elastic wave equation in heterogeneous media based on the energy method. Moreover, the plane wave analysis yields a sufficient and necessary stability condition by the von Neumann criterion in homogeneous case. Numerical computations for 3D wave simulation with point source excitation are given.

Keywords

3D, Elastic Wave, Inhomogeneous Media, Staggered-Grid Scheme, High-Order, Stability, Energy Estimate

1. Introduction

Numerical simulation of wave propagation has important applications in many scientific fields such as geophysics and seismic inversion. There are several types of numerical methods to solve the wave equations, for example, the finite difference method, the finite element method [1] [2] [3] [4], the spectral element method [5], the discontinuous Galerkin method [6] [7] and the finite volume method [8] [9]. Each of the above numerical methods has its own advantages and disadvantages. In this paper, we consider the finite difference method.

The finite difference method is a very popular method because of high computational efficiency. In fact, it has been applied to wave simulation for several decades [10] [11] [12] [13]. Since perfect numerical simulation depends on both stability and the order of accuracy, the high-order schemes and the corresponding

stability are an important research topic of this field. In particularly, we may list a few here. In [14], Cohen and Joly construct and analyses a family of fourth-order schemes for the acoustic wave equation in nonhomogeneous media. In [15], Sei analysis the stability of high-order difference schemes for the 2D elastic wave equation in heterogeneous media. The stable difference approximation for the 3D elastic wave equation in the second-order formulation in heterogeneous media has been investigated in [16]. In [17], a new family of locally one-dimensional schemes with fourth-order accuracy both in space and time for the 3D elastic wave equation is constructed and the stability is derived. The constructed new schemes in [17] only involve a three-point stencil in each spatial direction to achieve fourth-order accuracy. In this paper, based on the energy method, we study the stability analysis for the high-order staggered-grid schemes of the 3D elastic wave equation in heterogeneous media. To our knowledge, there is no work in this respect and our result is new.

The reminder of the paper is organized as follows. In Section 2, we present the governing equation and the high-order difference schemes in heterogeneous on staggered-grid grids. In Section 3, the stability analysis for the high-order difference schemes in heterogeneous is presented. In Section 4, the plane wave analysis in homogeneous case is investigated. In Section 5, we present numerical comparisons for 3D elastic wave simulation. Finally the conclusion and discussions are given in Section 6.

2. High-Order Spatial Discretization

We consider the following three-dimensional (3D) elastic wave equations in isotropic heterogeneous media

$$\begin{cases} \rho\dfrac{\partial^2 u}{\partial t^2}-\dfrac{\partial}{\partial x}\left((\lambda+2\mu)\dfrac{\partial u}{\partial x}+\lambda\dfrac{\partial v}{\partial y}+\lambda\dfrac{\partial w}{\partial z}\right)-\dfrac{\partial}{\partial y}\left(\mu\dfrac{\partial v}{\partial x}+\mu\dfrac{\partial u}{\partial y}\right)-\dfrac{\partial}{\partial z}\left(\mu\dfrac{\partial u}{\partial z}+\mu\dfrac{\partial w}{\partial x}\right)=f_1, \\ \rho\dfrac{\partial^2 v}{\partial t^2}-\dfrac{\partial}{\partial x}\left(\mu\dfrac{\partial v}{\partial x}+\mu\dfrac{\partial u}{\partial y}\right)-\dfrac{\partial}{\partial y}\left((\lambda+2\mu)\dfrac{\partial v}{\partial y}+\lambda\dfrac{\partial u}{\partial x}+\lambda\dfrac{\partial w}{\partial z}\right)-\dfrac{\partial}{\partial z}\left(\mu\dfrac{\partial v}{\partial z}+\mu\dfrac{\partial w}{\partial y}\right)=f_2, \\ \rho\dfrac{\partial^2 w}{\partial t^2}-\dfrac{\partial}{\partial x}\left(\mu\dfrac{\partial u}{\partial z}+\mu\dfrac{\partial w}{\partial x}\right)-\dfrac{\partial}{\partial y}\left(\mu\dfrac{\partial v}{\partial z}+\mu\dfrac{\partial w}{\partial y}\right)-\dfrac{\partial}{\partial z}\left((\lambda+2\mu)\dfrac{\partial w}{\partial z}+\lambda\dfrac{\partial u}{\partial x}+\lambda\dfrac{\partial v}{\partial y}\right)=f_3, \end{cases} \quad (1)$$

where $(u,v,w)(\boldsymbol{x},t)$ are the displacement vector at location $\boldsymbol{x}=(x,y,z)$ and time t, $\rho(\boldsymbol{x})$ is the density, $\lambda(\boldsymbol{x})>0$ and $\mu(\boldsymbol{x})\geq 0$ are the Lamé parameters, $\boldsymbol{f}=(f_1,f_2,f_3)$ is the external force.

Using the stress tensor, we can formulate the above system (1) as a first order in the following ways

$$\begin{cases} \rho\dfrac{\partial^2 u}{\partial t^2}-\left(\dfrac{\partial \tau^{xx}}{\partial x}+\dfrac{\partial \tau^{xy}}{\partial y}+\dfrac{\partial \tau^{xz}}{\partial z}\right)=0, \\ \rho\dfrac{\partial^2 v}{\partial t^2}-\left(\dfrac{\partial \tau^{xy}}{\partial x}+\dfrac{\partial \tau^{yy}}{\partial y}+\dfrac{\partial \tau^{yz}}{\partial z}\right)=0, \\ \rho\dfrac{\partial^2 w}{\partial t^2}-\left(\dfrac{\partial \tau^{xz}}{\partial x}+\dfrac{\partial \tau^{yz}}{\partial y}+\dfrac{\partial \tau^{zz}}{\partial z}\right)=0, \end{cases} \quad (2)$$

where

$$\begin{cases} \tau^{xx} = (\lambda + 2\mu)\dfrac{\partial u}{\partial x} + \lambda\dfrac{\partial v}{\partial y} + \lambda\dfrac{\partial w}{\partial z}, \\ \tau^{yy} = (\lambda + 2\mu)\dfrac{\partial v}{\partial y} + \lambda\dfrac{\partial u}{\partial x} + \lambda\dfrac{\partial w}{\partial z}, \\ \tau^{zz} = (\lambda + 2\mu)\dfrac{\partial w}{\partial z} + \lambda\dfrac{\partial u}{\partial x} + \lambda\dfrac{\partial v}{\partial y}, \\ \tau^{xy} = \mu\dfrac{\partial v}{\partial x} + \mu\dfrac{\partial u}{\partial y}, \\ \tau^{xz} = \mu\dfrac{\partial u}{\partial z} + \mu\dfrac{\partial w}{\partial x}, \\ \tau^{yz} = \mu\dfrac{\partial v}{\partial z} + \mu\dfrac{\partial w}{\partial y}. \end{cases} \tag{3}$$

Let Δx, Δy and Δz be the spatial steps of x, y, z directions respectively. Now discretization of (2) with the second-order accuracy in space gives

$$\rho\frac{\partial^2 u}{\partial t^2}(i,j,k) - \frac{\tau^{xx}_{i+\frac{1}{2},j,k} - \tau^{xx}_{i-\frac{1}{2},j,k}}{\Delta x} - \frac{\tau^{xy}_{i,j+\frac{1}{2},k} - \tau^{xy}_{i,j-\frac{1}{2},k}}{\Delta y} - \frac{\tau^{xz}_{i,j,k+\frac{1}{2}} - \tau^{xz}_{i,j,k-\frac{1}{2}}}{\Delta z} = 0, \tag{4}$$

$$\rho\frac{\partial^2 v}{\partial t^2}(i,j,k) - \frac{\tau^{xy}_{i+\frac{1}{2},j,k} - \tau^{xy}_{i-\frac{1}{2},j,k}}{\Delta x} - \frac{\tau^{yy}_{i,j+\frac{1}{2},k} - \tau^{yy}_{i,j-\frac{1}{2},k}}{\Delta y} - \frac{\tau^{yz}_{i,j,k+\frac{1}{2}} - \tau^{yz}_{i,j,k-\frac{1}{2}}}{\Delta z} = 0, \tag{5}$$

$$\rho\frac{\partial^2 w}{\partial t^2}(i,j,k) - \frac{\tau^{xz}_{i+\frac{1}{2},j,k} - \tau^{xz}_{i-\frac{1}{2},j,k}}{\Delta x} - \frac{\tau^{yz}_{i,j+\frac{1}{2},k} - \tau^{yz}_{i,j-\frac{1}{2},k}}{\Delta y} - \frac{\tau^{zz}_{i,j,k+\frac{1}{2}} - \tau^{zz}_{i,j,k-\frac{1}{2}}}{\Delta z} = 0. \tag{6}$$

For computing this, we need the values of u, v, and w at the grid $\left(i+\frac{1}{2}, j+\frac{1}{2}, k+\frac{1}{2}\right)$. One convenient way is to choose averaging the corresponding vales. For example,

$$u_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} = \frac{1}{3}\left(\frac{u_{i+1,j,k} + u_{i,j,k}}{2} + \frac{u_{i,j+1,k} + u_{i,j,k}}{2} + \frac{u_{i,j,k+1} + u_{i,j,k}}{2}\right).$$

However, such choices have no physical meaning. Another way is to compute u, v, w directly on staggered grids. In particular, we replace Equations (4)-(6) with Equations (7)-(9):

$$\rho\frac{\partial^2 u}{\partial t^2}(i,j,k) - \frac{\tau^{xx}_{i+\frac{1}{2},j,k} - \tau^{xx}_{i-\frac{1}{2},j,k}}{\Delta x} - \frac{\tau^{xy}_{i,j+\frac{1}{2},k} - \tau^{xy}_{i,j-\frac{1}{2},k}}{\Delta y} - \frac{\tau^{xz}_{i,j,k+\frac{1}{2}} - \tau^{xz}_{i,j,k-\frac{1}{2}}}{\Delta z} = 0, \tag{7}$$

$$\rho\frac{\partial^2 v}{\partial t^2}\left(i+\frac{1}{2}, j+\frac{1}{2}, k\right) - \frac{\tau^{xy}_{i+1,j+\frac{1}{2},k} - \tau^{xy}_{i,j+\frac{1}{2},k}}{\Delta x} - \frac{\tau^{yy}_{i+\frac{1}{2},j+1,k} - \tau^{yy}_{i+\frac{1}{2},j,k}}{\Delta y} - \frac{\tau^{yz}_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} - \tau^{yz}_{i+\frac{1}{2},j+\frac{1}{2},k-\frac{1}{2}}}{\Delta z} = 0, \tag{8}$$

$$\rho\frac{\partial^2 w}{\partial t^2}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)-\frac{\tau^{xz}_{i+1,j,k+\frac{1}{2}}-\tau^{xz}_{i,j,k+\frac{1}{2}}}{\Delta x}-\frac{\tau^{yz}_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}}-\tau^{yz}_{i+\frac{1}{2},j-\frac{1}{2},k+\frac{1}{2}}}{\Delta y}-\frac{\tau^{zz}_{i+\frac{1}{2},j,k+1}-\tau^{zz}_{i+\frac{1}{2},j,k}}{\Delta z}=0. \tag{9}$$

Obviously, the schemes (7)-(9) are the second-order accuracy in space. In order to construct high-order accuracy scheme in space, we first define the following functional spaces. Now we introduce the differentiation operator D_x on half integer grids with $O\left(\Delta x^{2L}\right)$ order as follows:

$$\left(\frac{\partial u}{\partial x}\right)_{i,j,k}\approx D_x u_{i,j,k}=\sum_{l=1}^{L}\frac{\beta_l}{\Delta x}\left[u_{i+l-\frac{1}{2},j,k}-u_{i-l+\frac{1}{2},j,k}\right], \tag{10}$$

or

$$\left(\frac{\partial u}{\partial x}\right)_{i+\frac{1}{2},j,k}\approx D_x u_{i+\frac{1}{2},j,k}=\sum_{l=1}^{L}\frac{\beta_l}{\Delta x}\left[u_{i+l,j,k}-u_{i-l+1,j,k}\right], \tag{11}$$

where β_l is difference coefficients on the staggered grids, which can be calculated by a fast algorithm [18] [19] [20] by Matlab tool. Obviously, the approximation (10) or (11) has $O\left(\Delta x^{2L}\right)$ order accuracy. For example, when $\beta_1=1$ for $L=1$ it has the second-order accuracy $O\left(\Delta x^2\right)$. And when $\beta_1=\frac{9}{8}$ and $\beta_2=\frac{1}{24}$ for $L=2$ it has the fourth-order accuracy $O\left(\Delta x^4\right)$. The general analytical expression of β_l is given in **Appendix**. Similarly, we can define the operators D_y and D_z. Here, the subscript of the operator refers to the direction of differentiation.

Now, we can construct the semi-discrete schemes of system (1)

$$\left(\rho\frac{\partial^2 u}{\partial t^2}+D_x\left((\lambda+2\mu)D_x u+\lambda D_y v+\lambda D_z w\right)+D_y\left(\mu D_x v+\mu D_y u\right)+D_z\left(\mu D_z u+\mu D_x w\right)\right)(i,j,k)=0, \tag{12}$$

$$\left(\rho\frac{\partial^2 v}{\partial t^2}+D_x\left(\mu D_x v+\mu D_y u\right)+D_y\left((\lambda+2\mu)D_y v+\lambda D_x u+\lambda D_z w\right)+D_z\left(\mu D_z v+\mu D_y w\right)\right)\left(i+\frac{1}{2},j+\frac{1}{2},k\right)=0, \tag{13}$$

$$\left(\rho\frac{\partial^2 w}{\partial t^2}+D_x\left(\mu D_z u+\mu D_x w\right)+D_y\left(\mu D_z v+\mu D_y w\right)+D_z\left((\lambda+2\mu)D_z w+\lambda D_x u+\lambda D_y v\right)\right)\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)=0. \tag{14}$$

Applying the central difference approximation for time with the second-order accuracy, we obtain the full-discrete schemes of system (1), we obtain

$$\rho_{i,j,k}\left(u^{n+1}-2u^{n}+u^{n-1}\right)_{i,j,k}+\Delta t^{2}\left\{D_{x}\left(\left(\lambda+2\mu\right)D_{x}u+\lambda D_{y}v+\lambda D_{z}w\right)\right.$$
$$\left.+D_{y}\left(\mu D_{x}v+\mu D_{y}u\right)+D_{z}\left(\mu D_{z}u+\mu D_{x}w\right)\right\}_{i,j,k}^{n}=0, \tag{15}$$

$$l\rho_{i+\frac{1}{2},j+\frac{1}{2},k}\left(v^{n+1}-2v^{n}+v^{n-1}\right)_{i+\frac{1}{2},j+\frac{1}{2},k}+\Delta t^{2}\left\{D_{x}\left(\mu D_{x}v+\mu D_{y}^{0}u\right)\right.$$
$$\left.+D_{y}\left(\left(\lambda+2\mu\right)D_{y}v+\lambda D_{x}u+\lambda D_{z}w\right)+D_{z}\left(\mu D_{z}v+\mu D_{y}w\right)\right\}_{i+\frac{1}{2},j+\frac{1}{2},k}^{n}=0, \tag{16}$$

$$\rho_{i+\frac{1}{2},j,k+\frac{1}{2}}\left(w^{n+1}-2w^{n}+w^{n-1}\right)_{i+\frac{1}{2},j,k+\frac{1}{2}}+\Delta t^{2}\left\{D_{x}\left(\mu D_{z}u+\mu D_{x}w\right)\right.$$
$$\left.+D_{y}\left(\mu D_{z}v+\mu D_{y}w\right)+D_{z}\left(\left(\lambda+2\mu\right)D_{z}w+\lambda D_{x}u+\lambda D_{y}v\right)\right\}_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n}=0, \tag{17}$$

where n denotes the time index and Δt the time step.

3. Stability Analysis

We now turn to the study of the numerical stability of the schemes (15)-(17). We are going to proceed by the energy method in analogy with continuous energy given by:

$$E=E_{c}+E_{p}, \tag{18}$$

with

$$E_{c}=\frac{1}{2}\int_{R^{3}}\rho\left[\left(\frac{\partial u}{\partial t}\right)^{2}+\left(\frac{\partial v}{\partial t}\right)^{2}+\left(\frac{\partial w}{\partial t}\right)^{2}\right]\mathrm{d}x\mathrm{d}y\mathrm{d}z,$$

$$E_{p}=\frac{1}{2}\int_{R^{3}}\lambda\left(\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}+\frac{\partial u}{\partial z}\right)^{2}+2\mu\left[\left(\frac{\partial u}{\partial x}\right)^{2}+\left(\frac{\partial v}{\partial y}\right)^{2}+\left(\frac{\partial w}{\partial z}\right)^{2}\right]$$
$$+\mu\left[\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)^{2}+\left(\frac{\partial u}{\partial z}+\frac{\partial w}{\partial x}\right)^{2}+\left(\frac{\partial v}{\partial z}+\frac{\partial w}{\partial y}\right)^{2}\right]\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

In a source-free infinite medium, the energy is conservative, *i.e.*, $\frac{\mathrm{d}E}{\mathrm{d}t}=0$. The discrete energy at time $\left(n+\frac{1}{2}\right)\Delta t$ is $E^{n+\frac{1}{2}}=E_{c}^{n+\frac{1}{2}}+E_{p}^{n+\frac{1}{2}}$. In order to compute $E_{c}^{n+\frac{1}{2}}$ and $E_{p}^{n+\frac{1}{2}}$ and analyze the stability, we define the following functional spaces

$$L_{000}^{3}=\left\{\varphi\in L^{3}\left(R^{3}\right)\,|\,\varphi=\sum_{i,j,k=-\infty}^{\infty}\varphi_{i,j,k}\right.$$
$$\left.\times I_{\left[\left(i-\frac{1}{2}\right)\Delta x,\left(i+\frac{1}{2}\right)\Delta x\right]\times\left[\left(j-\frac{1}{2}\right)\Delta y,\left(j+\frac{1}{2}\right)\Delta y\right]\times\left[\left(k-\frac{1}{2}\right)\Delta z,\left(k+\frac{1}{2}\right)\Delta z\right]}\left(x,y,z\right)\right\},$$

$$L_{***}^{3}=\left\{\varphi\in L^{3}\left(R^{3}\right)\,|\,\varphi=\sum_{i,j,k=-\infty}^{\infty}\varphi_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}}\right.$$
$$\left.\times I_{\left[i\Delta x,(i+1)\Delta x\right]\times\left[j\Delta y,(j+1)\Delta y\right]\times\left[k\Delta z,(k+1)\Delta z\right]}\left(x,y,z\right)\right\},$$

$$L^3_{0**}=\left\{\varphi\in L^3\left(R^3\right)\mid\varphi=\sum_{i,j,k=-\infty}^{\infty}\varphi_{i,j+\frac{1}{2},k+\frac{1}{2}}\times I_{\left[\left(i-\frac{1}{2}\right)\Delta x,\left(i+\frac{1}{2}\right)\Delta x\right]\times\left[j\Delta y,(j+1)\Delta y\right]\times\left[k\Delta z,(k+1)\Delta z\right]}(x,y,z)\right\},$$

where

$$I_{[a,b]\times[c,d]\times[e,f]}(x,y,z)=\begin{cases}1, & (x,y,z)\in[a,b]\times[c,d]\times[e,f],\\ 0, & (x,y,z)\notin[a,b]\times[c,d]\times[e,f],\end{cases}$$

where 0 represents the integer grid (i,j,k) and $*$ the half integer grid $i+\frac{1}{2}$ or $j+\frac{1}{2}$ or $k+\frac{1}{2}$. The other functional spaces L^3_{00*}, L^3_{**0}, L^3_{*00} L^3_{*00}, L^3_{*0*}, L^3_{0**} can be defined similarly. For saving space, we omit their definitions. Let the scalar inner product be defined in L^3_{000} by

$$(f,g,h)_{000}=\sum_{i,j,k=-\infty}^{\infty}f_{i,j,k}g_{i,j,k}h_{i,j,k}\Delta x\Delta y\Delta z.$$

Other inner products such as $(f,g,h)_{***}$, $(f,g,h)_{0**}$ and so on have similar meaning. In the following, we compute $E_c^{n+\frac{1}{2}}$ and $E_p^{n+\frac{1}{2}}$ respectively.

$$\begin{aligned}E_c^{n+\frac{1}{2}}=\frac{1}{2}\Bigg\{&\left(\rho\frac{u^{n+1}-u^n}{\Delta t},\frac{u^{n+1}-u^n}{\Delta t}\right)_{000}+\left(\rho\frac{v^{n+1}-v^n}{\Delta t},\frac{v^{n+1}-v^n}{\Delta t}\right)_{**0}\\&+\left(\rho\frac{w^{n+1}-w^n}{\Delta t},\frac{w^{n+1}-w^n}{\Delta t}\right)_{0*0}-\frac{\Delta t^2}{4}\Bigg[\Bigg(\lambda D_x\frac{u^{n+1}-u^n}{\Delta t}\\&+\lambda D_y\frac{v^{n+1}-v^n}{\Delta t}+\lambda D_z\frac{w^{n+1}-w^n}{\Delta t},D_x\frac{u^{n+1}-u^n}{\Delta t}+D_y\frac{v^{n+1}-v^n}{\Delta t}\\&+D_z\frac{w^{n+1}-w^n}{\Delta t}\Bigg)_{*00}+\Bigg(\mu D_x\frac{v^{n+1}-v^n}{\Delta t}+\mu D_y\frac{u^{n+1}-u^n}{\Delta t},\\&D_x\frac{v^{n+1}-v^n}{\Delta t}+D_y\frac{u^{n+1}-u^n}{\Delta t}\Bigg)_{0*0}+\Bigg(\mu D_z\frac{u^{n+1}-u^n}{\Delta t}\\&+\mu D_x\frac{w^{n+1}-w^n}{\Delta t},D_z\frac{u^{n+1}-u^n}{\Delta t}+D_x\frac{w^{n+1}-w^n}{\Delta t}\Bigg)_{00*}\\&+\left(\mu D_y\frac{w^{n+1}-w^n}{\Delta t}+\mu D_z\frac{v^{n+1}-v^n}{\Delta t},\mu D_y\frac{w^{n+1}-w^n}{\Delta t}+\mu D_z\frac{v^{n+1}-v^n}{\Delta t}\right)_{***}\\&+\left(2\mu D_x\frac{u^{n+1}-u^n}{\Delta t},D_x\frac{u^{n+1}-u^n}{\Delta t}\right)_{*00}+\Bigg(2\mu D_y\frac{v^{n+1}-v^n}{\Delta t},\\&D_y\frac{v^{n+1}-v^n}{\Delta t}\Bigg)_{*00}+\left(2\mu D_z\frac{w^{n+1}-w^n}{\Delta t},D_z\frac{w^{n+1}-w^n}{\Delta t}\right)_{*00}\Bigg]\Bigg\},\end{aligned}$$

$$\begin{aligned}E_p^{n+\frac{1}{2}}=\frac{1}{2}\Bigg\{&\Bigg(\lambda D_x\frac{u^{n+1}+u^n}{2}+\lambda D_y\frac{v^{n+1}+v^n}{2}+\lambda D_z\frac{w^{n+1}+w^n}{2},\\&D_x\frac{u^{n+1}+u^n}{2}+D_y\frac{v^{n+1}+v^n}{2}+D_z\frac{w^{n+1}+w^n}{2}\Bigg)_{*00}\end{aligned}$$

$$+\left(2\mu D_x\frac{u^{n+1}+u^n}{2},D_x\frac{u^{n+1}+u^n}{2}\right)_{*00}+\left(2\mu D_y\frac{v^{n+1}+v^n}{2},\right.$$
$$\left.D_y\frac{v^{n+1}+v^n}{2}\right)_{0*0}+\left(2\mu D_z\frac{w^{n+1}+w^n}{2},D_z\frac{w^{n+1}+w^n}{2}\right)_{*00}$$
$$+\left(\mu D_y\frac{u^{n+1}+u^n}{2}+\mu D_x\frac{v^{n+1}+v^n}{2},D_y\frac{u^{n+1}+u^n}{2}+D_x\frac{v^{n+1}+v^n}{2}\right)_{0*0}$$
$$+\left(\mu D_z\frac{u^{n+1}+u^n}{2}+\mu D_x\frac{w^{n+1}+w^n}{2},D_z\frac{u^{n+1}+u^n}{2}+D_x\frac{w^{n+1}+w^n}{2}\right)_{00*}$$
$$+\left(\mu D_z\frac{v^{n+1}+v^n}{2}+\mu D_y\frac{w^{n+1}+w^n}{2},D_z\frac{v^{n+1}+v^n}{2}+D_y\frac{w^{n+1}+w^n}{2}\right)_{***}\Bigg\}.$$

We have conservation of the discrete energy, that is: $\frac{E^{n+\frac{1}{2}}-E^{n-\frac{1}{2}}}{\Delta t}=0$. The stability of the scheme will be proven if the potential energy $E_p^{n+\frac{1}{2}}$ and the kinetic energy $E_c^{n+\frac{1}{2}}$ are positive. Since $E_p^{n+\frac{1}{2}}$ is obviously positive, we need to find out under what conditions $E_c^{n+\frac{1}{2}}$ is positive.

The problem can be reformulated as: $\forall u\in L^3_{000}$, $\forall v\in L^3_{**0}$ and $\forall w\in L^3_{*0*}$ with

$$I=\left(\lambda D_x u+\lambda D_y v+\lambda D_z w,D_x u+D_y v+D_z w\right)_{*00}$$
$$+\left(\mu D_x v+\mu D_y u,D_x v+D_y u\right)_{0*0}+\left(\mu D_z u+\mu D_x w,D_z u+D_x w\right)_{00*}$$
$$+\left(\mu D_y w+\mu D_z v,D_y w+D_z v\right)_{***}+\left(2\mu D_x u,D_x u\right)_{*00}$$
$$+\left(2\mu D_y v,D_y v\right)_{*00}+\left(2\mu D_z w,D_z w\right)_{*00}$$

we look for the corresponding conditions. Since

$$\frac{\Delta t^2}{4}I\le\left(\rho u,u\right)_{000}+\left(\rho v,v\right)_{**0}+\left(\rho w,w\right)_{*0*},\tag{19}$$

we can bound I as follows:

$$I\le 2\Big[\left(\lambda D_x u,D_x u\right)_{*00}+\left(\lambda D_y v,D_y v\right)_{*00}+\left(\lambda D_z w,D_z w\right)_{*00}$$
$$+\left(\mu D_x v,D_x v\right)_{0*0}+\left(\mu D_y u,D_y u\right)_{0*0}+\left(\mu D_x w,D_x w\right)_{00*}$$
$$+\left(\mu D_z u,D_z^0 u\right)_{00*}+\left(\mu D_z v,D_z v\right)_{***}+\left(\mu D_y w,D_y w\right)_{***}$$
$$+\left(\mu D_x u,D_x u\right)_{*00}+\left(\mu D_y v,D_y^* v\right)_{*00}+\left(\mu D_z w,D_z w\right)_{*00}\Big],$$

or

$$I\le 2\Bigg[\underbrace{\left(\left(\lambda+\mu\right)D_x u,D_x u\right)_{*00}+\left(\mu D_y u,D_y u\right)_{0*0}+\left(\mu D_z u,D_z u\right)_{00*}}_{I_1}$$
$$+\underbrace{\left(\left(\lambda+\mu\right)D_y v,D_y v\right)_{*00}+\left(\mu D_x v,D_x v\right)_{0*0}+\left(\mu D_z v,D_z v\right)_{***}}_{I_2}$$

$$+\underbrace{\left((\lambda+\mu)D_z w, D_z w\right)_{*00}+\left(\mu D_x w, D_x w\right)_{00*}+\left(\mu D_y w, D_y w\right)_{***}}_{I_3}\Bigg].$$

Obviously, $I \le 2(I_1+I_2+I_3)$. In the following we estimate I_1, I_2 and I_3 respectively. Since

$$\begin{aligned}I_1 &= \sum_{i,j,k}(\lambda+\mu)_{i+\frac{1}{2},j,k}\left[\sum_{l=1}^{L}\frac{\beta_l}{\Delta x}\left(u_{i+l,j,k}-u_{i-l+1,j,k}\right)\right]^2\Delta x\Delta y\Delta z\\ &+\sum_{i,j,k}\mu_{i,j+\frac{1}{2},k}\left[\sum_{l=1}^{L}\frac{\beta_l}{\Delta y}\left(u_{i,j+l,k}-u_{i,j-l+1,k}\right)\right]^2\Delta x\Delta y\Delta z\\ &+\sum_{i,j,k}\mu_{i,j,k+\frac{1}{2}}\left[\sum_{l=1}^{L}\frac{\beta_l}{\Delta z}\left(u_{i,j,k+l}-u_{i,j,k-l+1}\right)\right]^2\Delta x\Delta y\Delta z,\end{aligned}$$

we have the following estimates for I_1:

$$\begin{aligned}I_1 &\le 2\left(\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta x}\right|\right)\sum_{i,j,k}\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta x}\right|(\lambda+\mu)_{i+\frac{1}{2},j,k}\left(u_{i+l,j,k}^2+u_{i-l+1,j,k}^2\right)\Delta x\Delta y\Delta z\\ &+2\left(\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta y}\right|\right)\sum_{i,j,k}\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta y}\right|\mu_{i,j,k+\frac{1}{2}}\left(u_{i,j,k+l}^2+u_{i,j,k-l+1}^2\right)\Delta x\Delta y\Delta z\\ &+2\left(\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta z}\right|\right)\sum_{i,j,k}\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta z}\right|\mu_{i,j+\frac{1}{2},k}\left(u_{i,j+l,k}^2+u_{i,j-l+1,k}^2\right)\Delta x\Delta y\Delta z,\end{aligned}$$

or

$$\begin{aligned}I_1 &\le 4\left(\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta x}\right|\right)\sum_{i,j,k}\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta x}\right|\left[\frac{(\lambda+\mu)_{i+l-\frac{1}{2},j,k}}{2\rho_{i,j,k}}+\frac{(\lambda+\mu)_{i-l+\frac{1}{2},j,k}}{2\rho_{i,j,k}}\right]\rho_{i,j,k}u_{i,j,k}^2\Delta x\Delta y\Delta z\\ &+4\left(\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta z}\right|\right)\sum_{i,j,k}\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta z}\right|\left[\frac{\mu_{i,j,k+l-\frac{1}{2}}}{2\rho_{i,j,k}}+\frac{\mu_{i,j,k-l+\frac{1}{2}}}{2\rho_{i,j,k}}\right]\rho_{i,j,k}u_{i,j,k}^2\Delta x\Delta y\Delta z\\ &+4\left(\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta y}\right|\right)\sum_{i,j,k}\sum_{l=1}^{L}\left|\frac{\beta_l}{\Delta y}\right|\left[\frac{\mu_{i,j+l-\frac{1}{2},k}}{2\rho_{i,j,k}}+\frac{\mu_{i,j-l+\frac{1}{2},k}}{2\rho_{i,j,k}}\right]\rho_{i,j,k}u_{i,j,k}^2\Delta x\Delta y\Delta z,\end{aligned}$$

or

$$\begin{aligned}I_1 &\le \frac{4}{\Delta x^2}\left(\sum_{l=1}^{L}|\beta_l|\right)\sum_{i,j,k}\sum_{l=1}^{L}|\beta_l|\left[\frac{(\lambda+\mu)_{i+l-\frac{1}{2},j,k}}{2\rho_{i,j,k}}+\frac{(\lambda+\mu)_{i-l+\frac{1}{2},j,k}}{2\rho_{i,j,k}}\right](\rho u,u)_{000}\\ &+\frac{4}{\Delta z^2}\left(\sum_{l=1}^{L}|\beta_l|\right)\sum_{i,j,k}\sum_{l=1}^{L}|\beta_l|\left[\frac{\mu_{i,j,k+l-\frac{1}{2}}}{2\rho_{i,j,k}}+\frac{\mu_{i,j,k-l+\frac{1}{2}}}{2\rho_{i,j,k}}\right](\rho u,u)_{000}\\ &+\frac{4}{\Delta y^2}\left(\sum_{l=1}^{L}|\beta_l|\right)\sum_{i,j,k}\sum_{l=1}^{L}|\beta_l|\left[\frac{\mu_{i,j+l-\frac{1}{2},k}}{2\rho_{i,j,k}}+\frac{\mu_{i,j-l+\frac{1}{2},k}}{2\rho_{i,j,k}}\right](\rho u,u)_{000}.\end{aligned}$$

Thus we obtain

$$I_1 \le \left(\frac{4}{\Delta x^2}+\frac{4}{\Delta y^2}+\frac{4}{\Delta z^2}\right)\left(\sum_{l=1}^{L}|\beta_l|\right)\max_{i,j,k}\sum_{l=1}^{L}|\beta_l|\times\left[\frac{(\lambda+\mu)_{i+l-\frac{1}{2},j,k}}{2\rho_{i,j,k}}\right.$$
$$\left.+\frac{(\lambda+\mu)_{i-l+\frac{1}{2},j,k}}{2\rho_{i,j,k}}+\frac{\mu_{i,j,k+l-\frac{1}{2}}}{2\rho_{i,j,k}}+\frac{\mu_{i,j,k-l+\frac{1}{2}}}{2\rho_{i,j,k}}+\frac{\mu_{i,j+l-\frac{1}{2},k}}{2\rho_{i,j,k}}+\frac{\mu_{i,j-l+\frac{1}{2},k}}{2\rho_{i,j,k}}\right](\rho u,u)_{000}.$$

Set

$$c_1^2=\left(\sum_{l=1}^{L}|\beta_l|\right)^{-1}\max_{i,j,k}\sum_{l=1}^{L}|\beta_l|\left[\frac{(\lambda+\mu)_{i+l-\frac{1}{2},j,k}}{2\rho_{i,j,k}}+\frac{(\lambda+\mu)_{i-l+\frac{1}{2},j,k}}{2\rho_{i,j,k}}\right.$$
$$\left.+\frac{\mu_{i,j,k+l-\frac{1}{2}}}{2\rho_{i,j,k}}+\frac{\mu_{i,j,k-l+\frac{1}{2}}}{2\rho_{i,j,k}}+\frac{\mu_{i,j+l-\frac{1}{2},k}}{2\rho_{i,j,k}}+\frac{\mu_{i,j-l+\frac{1}{2},k}}{2\rho_{i,j,k}}\right]. \tag{20}$$

Then for $\forall u\in L^3_{000}$, we have

$$I_1 \le \left(\frac{4}{\Delta x^2}+\frac{4}{\Delta y^2}+\frac{4}{\Delta z^2}\right)\left(\sum_{l=1}^{L}|\beta_l|\right)^2\cdot c_1^2\cdot(\rho u,u)_{000},\ \ \forall u\in L^3_{000}. \tag{21}$$

Similarly, we have

$$I_2 \le \left(\frac{4}{\Delta x^2}+\frac{4}{\Delta y^2}+\frac{4}{\Delta z^2}\right)\left(\sum_{l=1}^{L}|\beta_l|\right)^2\cdot c_2^2\cdot(\rho v,v)_{**0},\ \ \forall v\in L^3_{**0}, \tag{22}$$

where

$$c_2^2=\left(\sum_{l=1}^{L}|\beta_l|\right)^{-1}\max_{i,j,k}\sum_{l=1}^{L}|\beta_l|\left[\frac{(\lambda+\mu)_{i+\frac{1}{2},j+l,k}}{2\rho_{i+\frac{1}{2},j+\frac{1}{2},k}}+\frac{(\lambda+\mu)_{i+\frac{1}{2},j-l,k}}{2\rho_{i+\frac{1}{2},j+\frac{1}{2},k}}\right.$$
$$\left.+\frac{\mu_{i+l,j+\frac{1}{2},k}}{2\rho_{i+\frac{1}{2},j+\frac{1}{2},k}}+\frac{\mu_{i-l,j+\frac{1}{2},k}}{2\rho_{i+\frac{1}{2},j+\frac{1}{2},k}}+\frac{\mu_{i+\frac{1}{2},j+\frac{1}{2},k+l-\frac{1}{2}}}{2\rho_{i+\frac{1}{2},j+\frac{1}{2},k}}+\frac{\mu_{i+\frac{1}{2},j+\frac{1}{2},k-l+\frac{1}{2}}}{2\rho_{i+\frac{1}{2},j+\frac{1}{2},k}}\right]. \tag{23}$$

And

$$I_3 \le \left(\frac{4}{\Delta x^2}+\frac{4}{\Delta y^2}+\frac{4}{\Delta z^2}\right)\left(\sum_{l=1}^{L}|\beta_l|\right)^2\cdot c_3^2\cdot(\rho w,w)_{*0*},\ \ \forall w\in L^3_{*0*}, \tag{24}$$

where

$$c_3^2=\left(\sum_{l=1}^{L}|\beta_l|\right)^{-1}\max_{i,j,k}\sum_{l=1}^{L}|\beta_l|\left[\frac{(\lambda+\mu)_{i+\frac{1}{2},j,k+l}}{2\rho_{i+\frac{1}{2},j,k+\frac{1}{2}}}+\frac{(\lambda+\mu)_{i+\frac{1}{2},j,k-l+1}}{2\rho_{i+\frac{1}{2},j,k+\frac{1}{2}}}\right.$$
$$\left.+\frac{\mu_{i+l,j,k+\frac{1}{2}}}{2\rho_{i+\frac{1}{2},j,k+\frac{1}{2}}}+\frac{\mu_{i-l+1,j+\frac{1}{2},k+\frac{1}{2}}}{2\rho_{i+\frac{1}{2},j,k+\frac{1}{2}}}+\frac{\mu_{i+\frac{1}{2},j+l-\frac{1}{2},k+\frac{1}{2}}}{2\rho_{i+\frac{1}{2},j,k+\frac{1}{2}}}+\frac{\mu_{i+\frac{1}{2},j-l+\frac{1}{2},k+\frac{1}{2}}}{2\rho_{i+\frac{1}{2},j,k+\frac{1}{2}}}\right]. \tag{25}$$

Substituting (21), (22) and (24) into (19), we have

$$c^2\Delta t^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}+\frac{1}{\Delta z^2}\right)\left(\sum_{l=1}^{L}|\beta_l|\right)^2 \le \frac{1}{2}. \tag{26}$$

where $c=\max\{c_1,c_2,c_3\}$. Thus we obtain the sufficient stability condition for the numerical scheme (15)-(17). If the grid is uniform, *i.e.* $\Delta x=\Delta y=\Delta z\equiv h$, then (26) gives

$$c\Delta t \le \frac{\sqrt{6}}{6}h\left(\sum_{l=1}^{L}|\beta_l|\right)^{-1}.$$

Therefore we summarize the conclusion above into the following theorem.

Theorem 1. A sufficient stability condition for the numerical schemes (15)-(17) is

$$c\Delta t\sqrt{\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}+\frac{1}{\Delta z^2}} \le \frac{\sqrt{2}}{2}\left(\sum_{l=1}^{L}|\beta_l|\right)^{-1}, \tag{27}$$

If $\Delta x=\Delta y=\Delta z\equiv h$, then it reduces to

$$c\Delta t \le \frac{\sqrt{6}}{6}h\left(\sum_{l=1}^{L}|\beta_l|\right)^{-1}, \tag{28}$$

where $c=\max\{c_1,c_2,c_3\}$, and c_1, c_2 and c_3 are given by (20), (23) and (25) respectively.

4. Plane Wave Analysis

We turn to Fourier analysis [21] and we will derive the dispersion relation and by the von Neumann criterion we will get a necessary and sufficient stability condition. In homogeneous case for (12)-(14), the full-discrete schemes can be written as

$$\begin{aligned}&\rho\left(u^{n+1}-2u^n+u^{n-1}\right)_{i,j,k}+\Delta t^2\left\{(\lambda+2\mu)D_x^2u^n+\lambda D_xD_yv^n+\lambda D_xD_zw^n\right.\\&\left.+\mu D_yD_xv^n+\mu D_y^2u^n+\mu D_z^2u^n+\mu D_zD_xw^n\right\}(i,j,k)=0,\end{aligned} \tag{29}$$

$$\begin{aligned}&\rho\left(v^{n+1}-2v^n+v^{n-1}\right)_{i+\frac{1}{2},j+\frac{1}{2},k}+\Delta t^2\left\{\mu D_x^2v^n+\mu D_xD_yu^n+(\lambda+2\mu)D_y^2v^n\right.\\&\left.+\lambda D_yD_xu^n+\lambda D_yD_zw^n+\mu D_z^2v^n+\mu D_zD_yw^n\right\}\left(i+\frac{1}{2},j+\frac{1}{2},k\right)=0,\end{aligned} \tag{30}$$

$$\begin{aligned}&\rho\left(w^{n+1}-2w^n+w^{n-1}\right)_{i+\frac{1}{2},j,k+\frac{1}{2}}+\Delta t^2\left\{\mu D_xD_zu^n+\mu D_x^2w^n+\mu D_yD_zv^n\right.\\&\left.+\mu D_y^2w^n+(\lambda+2\mu)D_z^2w^n+\lambda D_zD_xu^n+\lambda D_zD_yv^n\right\}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right)=0.\end{aligned} \tag{31}$$

We assume that $\boldsymbol{u}=\boldsymbol{d}\mathrm{e}^{\mathrm{i}(\omega t-\boldsymbol{k}\cdot\boldsymbol{x})}$ is a solution of Equation (29)-(31), where $\mathrm{i}=\sqrt{-1}$, ω is the angular frequency, $\boldsymbol{d}=(d_1,d_2,d_3)$ amplitude, and

$$\boldsymbol{k}=(k_1,k_2,k_3)\equiv|\boldsymbol{k}|(\cos\theta\sin\phi,\sin\theta\sin\phi,\cos\phi)$$

is the wave vector. Here θ is the propagation angle and ϕ the propagation azimuth. The two angles determine the movement direction of the plane wave in the 3D space.

Substituting the plane wave solution into Equations (29)-(31), we obtain the

following relations:

$$
\begin{aligned}
& d_1 \sin^2\left(\frac{\omega\Delta t}{2}\right) \\
& = \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+2\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\right)^2\right. \\
& \left. + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\right)^2 + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right)^2\right\} d_1 \qquad (32) \\
& + \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\right\} d_2 \\
& + \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right\} d_3,
\end{aligned}
$$

$$
\begin{aligned}
& d_2 \sin^2\left(\frac{\omega\Delta t}{2}\right) \\
& = \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+2\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\right)^2\right. \\
& \left. + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{1}\right)\right)^2 + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right)^2\right\} d_2 \qquad (33) \\
& + \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\right\} d_3 \\
& + \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right\} d_1,
\end{aligned}
$$

$$
\begin{aligned}
& d_3 \sin^2\left(\frac{\omega\Delta t}{2}\right) \\
& = \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+2\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right)^2\right. \\
& \left. + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\right)^2 + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\right)^2\right\} d_3 \qquad (34) \\
& + \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right\} d_1 \\
& + \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right\} d_2.
\end{aligned}
$$

By introducing the matrix B with elements $\left(b_{ij}\right)$ defined by

$$
\begin{aligned}
b_{11} &= \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+2\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\right)^2\right. \\
&\quad \left. + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\right)^2 + \frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_3 h}{2}\right)\right)^2\right\}, \\
b_{12} &= b_{21} = \frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_1 h}{2}\right)\sum_{l=1}^{L}\beta_l \sin\left((2l-1)\frac{k_2 h}{2}\right)\right\},
\end{aligned}
$$

$$b_{22}=\frac{\Delta t^2}{h^2}\left\{\frac{\lambda+2\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_2h}{2}\right)\right)^2+\frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_1h}{2}\right)\right)^2+\frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_3h}{2}\right)\right)^2\right\},$$

$$b_{13}=b_{31}=\frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_1h}{2}\right)\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_3h}{2}\right)\right\},$$

$$b_{23}=b_{32}=\frac{\Delta t^2}{h^2}\left\{\frac{\lambda+\mu}{\rho}\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_2h}{2}\right)\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_3h}{2}\right)\right\},$$

$$b_{33}=\frac{\Delta t^2}{h^2}\left\{\frac{\lambda+2\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_3h}{2}\right)\right)^2+\frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_2h}{2}\right)\right)^2+\frac{\mu}{\rho}\left(\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{k_1h}{2}\right)\right)^2\right\},$$

we can write the relations (32)-(34) as the following matrix form

$$B\boldsymbol{d}=\sin^2\left(\frac{\omega\Delta t}{2}\right)\boldsymbol{d}. \tag{35}$$

The eigenvalues of B then express ω as a function of $\boldsymbol{k}$, which is the dispersion relation. There are three eigenvalues for matrix B. One eigenvalue is corresponding to the longitudinal or compressional wave, the double eigenvalues are corresponding to the transverse or shear wave. Thus we have the following two different relations

$$\begin{cases}\sin\left(\frac{\omega\Delta t}{2}\right)=\frac{C_p\Delta t}{h}\sqrt{A^2(k_1)+A^2(k_2)+A^2(k_3)},\\ \sin\left(\frac{\omega\Delta t}{2}\right)=\frac{C_s\Delta t}{h}\sqrt{A^2(k_1)+A^2(k_2)+A^2(k_3)},\end{cases} \tag{36}$$

where

$$A(k)=\sum_{l=1}^{L}\beta_l\sin\left((2l-1)\frac{kh}{2}\right),\ C_p=\sqrt{\frac{\lambda+2\mu}{\rho}},\ C_s=\sqrt{\frac{\mu}{\rho}}.$$

here C_p and C_s are the velocities of compressional and shear waves. Note that C_p is always larger than C_s. With the dispersion relations (36), we can apply the von Neumann stability criterion. A necessary stability is that the eigenvalues of B must be lower than 1. Thus we have

$$\frac{C_p\cdot\Delta t}{h}\sqrt{A^2(k_1)+A^2(k_2)+A^2(k_3)}\le 1. \tag{37}$$

It is easy to verify that

$$\max_{k_1,k_2,k_3\in\mathbb{R}}\sqrt{A^2(k_1)+A^2(k_2)+A^2(k_3)}=\sqrt{A^2\left(\frac{\pi}{h}\right)+A^2\left(\frac{\pi}{h}\right)+A^2\left(\frac{\pi}{h}\right)}=\sqrt{3}\left(\sum_{l=1}^{L}|\beta_l|\right). \tag{38}$$

Therefore we obtain the following theorem.

Theorem 2. In the homogeneous case, a sufficient and necessary stability condition for the numerical schemes (29)-(31) is given by

$$\Delta t \le \frac{\sqrt{3}}{3}\left(\sum_{l=1}^{L}|\beta_l|\right)^{-1}. \tag{39}$$

Proof Combining Equations (37) and (38), we obtain (39). Moreover, the matrix B is symmetric in homogeneous case. So the condition (39) is a sufficient and necessary condition. The proof is complete.

We now define the normalized phase error E_ϕ as follow:

$$E_\phi = \frac{C_\phi - C}{C_v},\ C_\phi = \frac{\omega(k)}{k}, \tag{40}$$

which is a function of $H \equiv \frac{|\boldsymbol{k}|h}{2\pi}$, where C_v indicates C_p or C_s which is related to different kinds of compressional wave and shear wave.

The stability condition $P \equiv C\Delta t/h$ is defined by Courant-Friedrichs-Lewy (CFL) condition which bounds the interval for stability. We plot some dispersion curves based on Equation (40). Without loss of generality, we present dispersion curves for some special propagation angle and azimuth. Figure 1 is the normalized phase error for fixed $\theta = 45^\circ$ and $\phi = 45^\circ$ with different values of CFL condition and it shows that the phase error drops as increasing the order of accuracy. Figure 2 shows the normalized phase error for $\theta = 30^\circ$ and different values of ϕ for different order or *L*. The figures for other propagation angle θ and azimuth ϕ are similar we omit them for saving space.

5. Numerical Computations

Wave simulation ignited by a point source is usually adopted in geophysical applications. For convenience, we simulate 3D elastic wave propagation in a homogeneous cubic model. The computational domain is $[0, 2000\ \mathrm{m}]^3$. The source is located in the center of the model and its time function is given by

$$s(t) = \sin(300t)\mathrm{e}^{-(300t)^3}, \tag{41}$$

which is loaded on the u component. The compressional velocity is 4000 m/s and the shear velocity 2500 m/s. The time step is $\Delta t = 0.001\,\mathrm{s}$ and the space step is $h = 10\ \mathrm{m}$. Figure 3 shows the 3D snapshot of u component at propagation time 0.2 s. For brevity, we present some 2D slices of the 3D snapshots of *u*, *v*, and *w* components. The *xz* sections of 3D snapshots of *u*, *v*, and *w* components at propagation time 0.2 s are shown in Figures 4-6 respectively. We omit other sections for space. In our computations the scheme with fourth-order accuracy in space is applied. We remark that the comparisons between the numerical solution and the exact solution can be found in [17]. From Figures 4-6, we can clearly see the two types of waves, *i.e.* the compressional wave and the shear wave, which is consistent with the physical phenomenon.

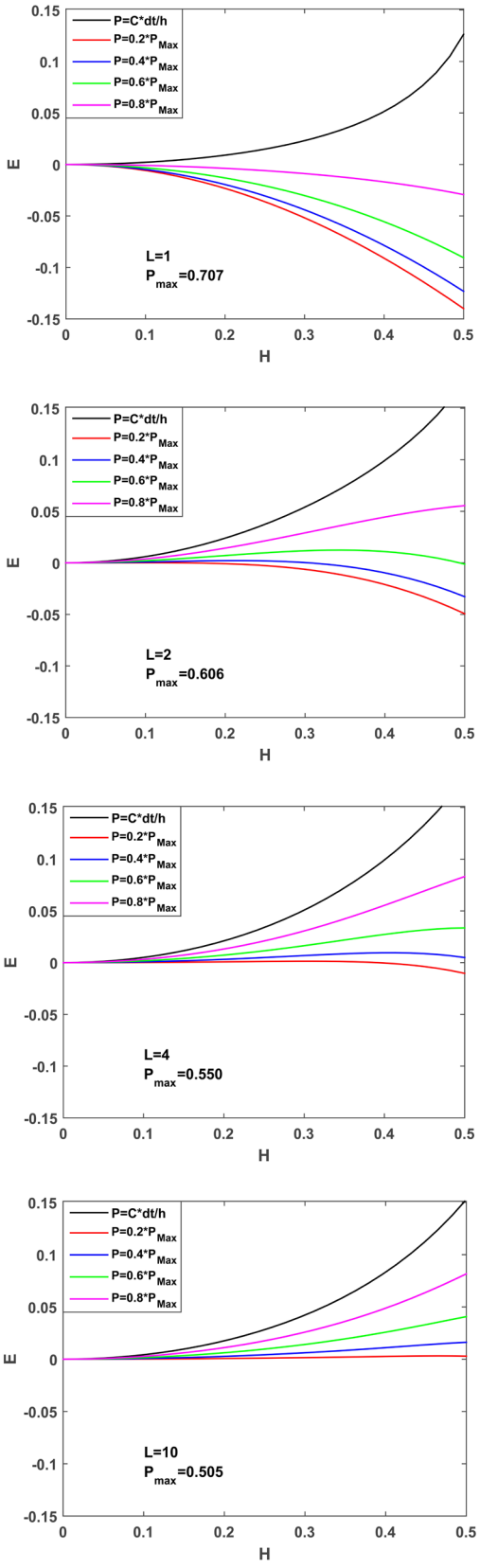


Figure 1. The normalized phase error for different CFL number at the stability limit $P_{\max} = C\Delta t/h$. The propagation angles $\theta = 45^{\circ}$ and $\phi = 45^{\circ}$ are fixed.

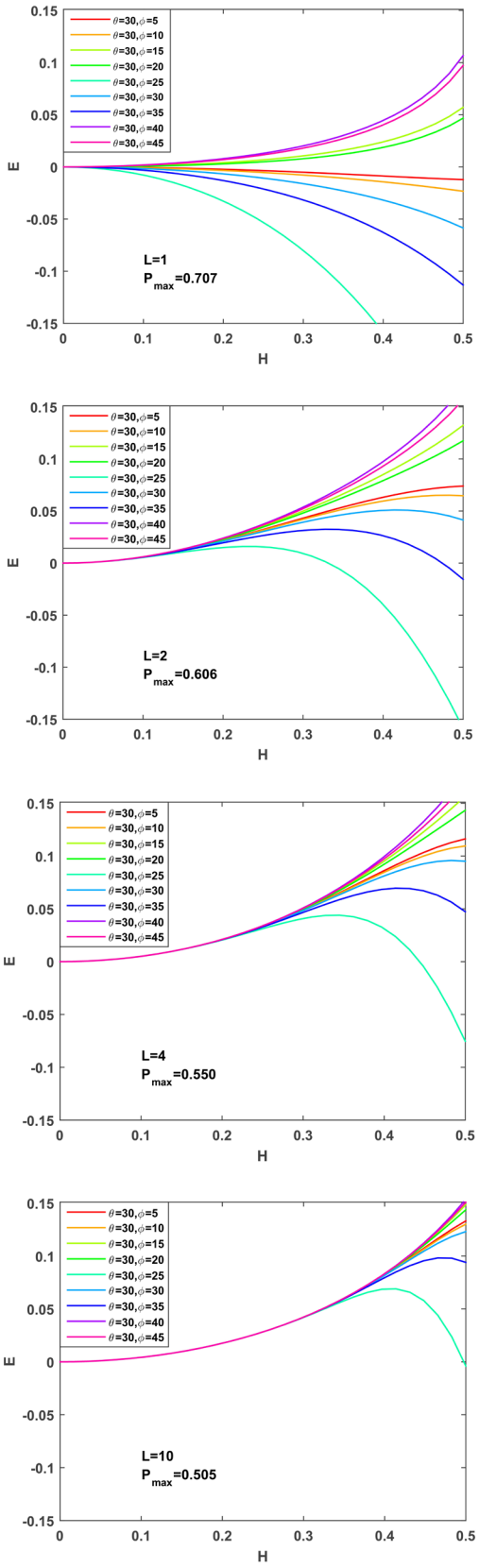


Figure 2. The normalized phase error for different CFL number at different propagation angle ϕ. The propagation angle $\theta = 30^{\circ}$ is fixed.

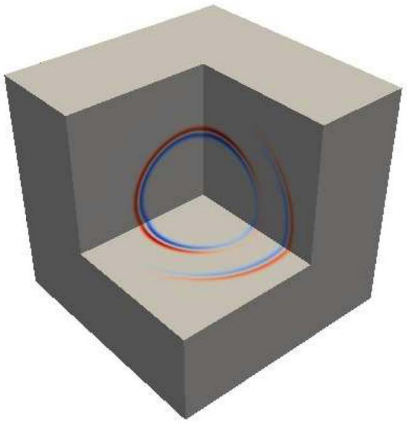


Figure 3. The 3D snapshot of u component at propagation time 0.2 s.

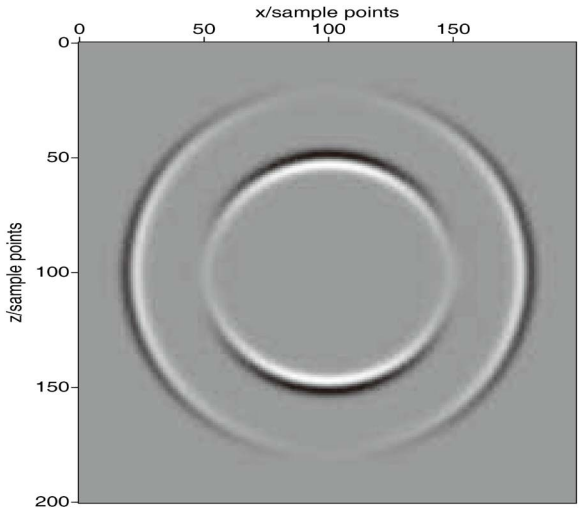
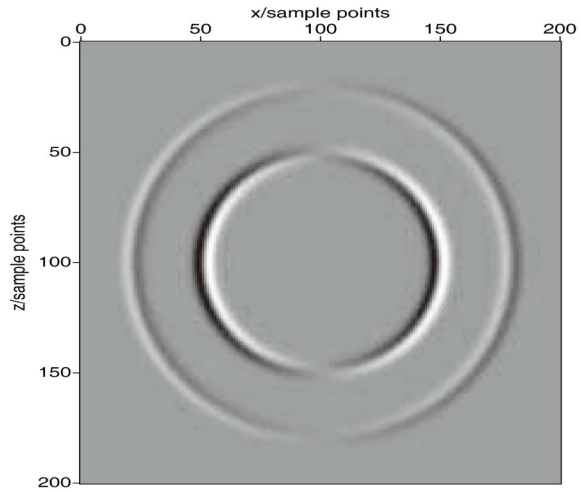


Figure 4. The xz-section of 3D snapshot of u component at propagation time 0.2 s.



Figures 5. The xz-section of 3D snapshot of v component at propagation time 0.2 s.

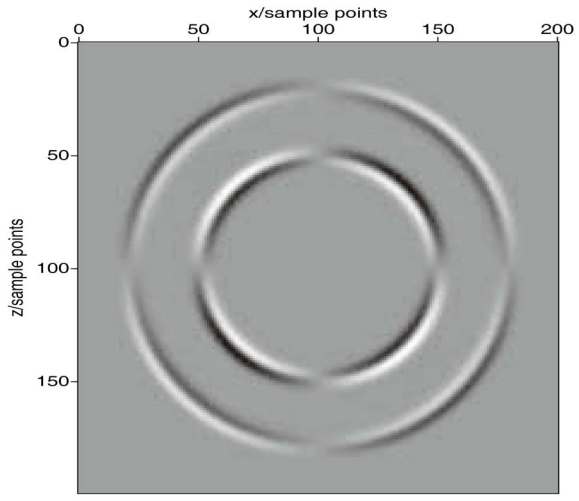


Figure 6. The *xz*-section of 3D snapshot of *w* component at propagation time 0.2 s.

6. Conclusion

The staggered-grid difference method is a very important technique to solve wave equations numerically because of its high efficiency and the character of energy preservation. It has been well applied to seismic wave propagation for more than two decades. Based on the energy estimate method, we implement the stability analysis for the high-order staggered-grid schemes of the inhomogeneous 3D elastic wave equation. The stability result is controlled by the space varying parameters and the difference coefficients. The plane wave analysis in homogeneous media is completed and by the von Neumann criterion a necessary and sufficient stability condition is obtained. The analysis is helpful to design the computational parameters such as the time step and the space steps. Numerical computations are given to verify the effectiveness of the schemes. The key point of this paper is the theoretical analysis. In the future, we will consider more numerical computations for inhomogeneous media.

Acknowledgements

This work is supported by National Natural Science Foundation of China under grant numbers 11471328 and 51739007. It is also partially supported by National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences.

Conflicts of Interest

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

References

[1] De Basabe, J.D. and Sen, M.K. (2010) Stability of the High-Order Finite Elements for Acoustic or Elastic Wave Propagation with High-Order Time Stepping. *Geo-*

physical Journal International, **181**, 577-590.
https://doi.org/10.1111/j.1365-246X.2010.04536.x

[2] Bécache, E., Joly, P. and Tsogka, C. (2002) A New Family of Mixed Finite Elements for the Linear Elastodynamic Problem. *SIAM Journal on Numerical Analysis*, **39**, 2109-2132. https://doi.org/10.1137/S0036142999359189

[3] Cohen, G.C. (2002) Higher-Order Numerical Methods for Transient Wave Equations. Springer, New York. https://doi.org/10.1007/978-3-662-04823-8

[4] Zhang, W., Chung, E.T. and Wang, C. (2014) Stability for Imposing Absorbing Boundary Conditions in the Finite Element Simulation of Acoustic Wave Propagation. *Journal of Computational Mathematics*, **32**, 1-20.
https://doi.org/10.4208/jcm.1310-m3942

[5] Komatitsch, D., Martin, R., Tromp, J., Taylor, M.A. and Wingate, B.A. (2001) Wave Propagation in 2-D Elastic Media Using a Spectral Element Method with Triangles and Quadrangles. *Journal of Computational Acoustics*, **9**, 703-718.
https://doi.org/10.1142/S0218396X01000796

[6] Chung, E.T. and Engquist, B. (2006) Optimal Discontinuous Galerkin Methods for Wave Propagation. *SIAM Journal on Numerical Analysis*, **44**, 2131-2158.
https://doi.org/10.1137/050641193

[7] Dumbser, M., Käser, M. and Toto, E.F. (2007) An Arbitrary High-Order Discontinuous Galerkin Method for Elastic Waves on Unstructured Meshes-V. Local Time Stepping and *p*-Adaptivity. *Geophysical Journal International*, **171**, 695-717.
https://doi.org/10.1111/j.1365-246X.2007.03427.x

[8] Dumbser, M., Käser, M. and de la Puente, J. (2007) Arbitrary High-Order Finite Volume Schemes for Seismic Wave Propagation on Unstructured Meshes in 2D and 3D. *Geophysical Journal International*, **171**, 665-694.
https://doi.org/10.1111/j.1365-246X.2007.03421.x

[9] Zhang, W., Zhuang, Y. and Chung, E.T. (2007) A New Spectral Finite Volume Method for Elastic Wave Modelling on Unstructured Meshes. *Geophysical Journal International*, **206**, 292-307. https://doi.org/10.1093/gji/ggw148

[10] Alford, R.M., Kelly, K.R. and Boore, D.M. (1974) Accuracy of Finite-Difference Modeling of the Acoustic Wave Equation. *Geophysics*, **39**, 834-842.
https://doi.org/10.1190/1.1440470

[11] Bayliss, A., Jordan, K.E., Lemesurier, B. and Turkel, E. (1986) A Fourth Accurate Finite Difference Scheme for the Computation of Elastic Waves. *Bulletin of the Seismological Society of America*, **76**, 1115-1132.

[12] Virieux, J. (1986) P-SV Wave Propagation in Heterogeneous Media: Velocity Stress Formulation Finite Difference Method. *Geophysics*, **51**, 889-901.
https://doi.org/10.1190/1.1441605

[13] Zingg, D.W., Lomax, H. and Jurgens, H. (1996) High-Accuracy Finite-Difference Schemes for Linear Wave Propagation. *SIAM Journal on Scientific Computing*, **17**, 328-346. https://doi.org/10.1137/S1064827599350320

[14] Cohen, G. and Joly, P. (1996) Construction and Analysis of Fourth-Order Finite Difference Schemes for the Acoustic Wave Equation in Nonhomogeneous Media. *SIAM Journal on Numerical Analysis*, **33**, 1266-1302.
https://doi.org/10.1137/S0036142993246445

[15] Sei, A. (1995) A Family of Numerical Schemes for the Computation of Elastic Waves. *SIAM Journal on Scientific Computing*, **16**, 898-916.
https://doi.org/10.1137/0916052

[16] Nilsson, S., Petersson, N.A., Sjögreen, B. and Kreiss, H.-O. (2007) Stable Difference Approximations for the Elastic Wave Equation in Second Order Formulation. *SIAM Journal on Numerical Analysis*, **45**, 1902-1936. https://doi.org/10.1137/060663520

[17] Zhang, W. (2019) A New Family of Fourth-Order Locally One-Dimensional Schemes for the 3D Elastic Wave Equation. *Journal of Computational and Applied Mathematics*, **348**, 246-260. https://doi.org/10.1016/j.cam.2018.08.056

[18] Fornberg, B. (1988) Generation of Finite Difference Formulas on Arbitrarily Spaced Grids. *Mathematics of Computation*, **51**, 699-706. https://doi.org/10.1090/S0025-5718-1988-0935077-0

[19] Fornberg, B. (1990) High-Order Finite Differences and the Pseudospectral Method on Staggered Grids. *SIAM Journal on Numerical Analysis*, **27**, 904-918. https://doi.org/10.1137/0727052

[20] Fornberg, B. and Ghrist, M. (1999) Spatial Finite Difference Approximations for Wave-Type Equations. *SIAM Journal on Numerical Analysis*, **37**, 105-130. https://doi.org/10.1137/S0036142998335881

[21] Thomas, J.W. (1995) Numerical Partial Differential Equations: Finite Difference Methods. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4899-7278-1

Appendix: Expression of the Coefficient β_l

The calculation of difference coefficients on both regular and staggered-grid grids has been investigated by several authors [15] [18] [19] [20]. In this appendix we present the analytical expression of the difference coefficient β_l in (10) or (11). In order to calculate the coefficient β_l, we consider an explicit staggered-grid difference expression for a function $f(x)$:

$$\frac{\partial f}{\partial x} \approx \frac{1}{h}\sum_{l=1}^{L}\beta_l\left[f\left(x+lh-\frac{1}{2}h\right)-f\left(x-lh+\frac{1}{2}h\right)\right],$$

where h is the step size, L is a positive number and β_l are the difference coefficients. Now consider $f = f_0 \mathrm{e}^{\mathrm{i}kx}$ and $\alpha = kh/2$, where k is the wave number, $\mathrm{i} = \sqrt{-1}$ and f_0 is a constant then we have

$$\alpha \approx \sum_{l=1}^{L}\beta_l \sin\left((2l-1)\alpha\right).$$

Then Taylor's series expansion gives

$$\alpha \approx \sum_{m=1}^{\infty}\left[\frac{(-1)^{m-1}}{(2m-1)!}\sum_{l=1}^{L}(2l-1)^{2m-1}\beta_l\alpha^{2m-1}\right].$$

Now equating the coefficient of α both sides we get

$$\frac{(-1)^{m-1}}{(2m-1)!}\sum_{l=1}^{L}\left[(2l-1)^{2m-1}\right]\beta_l = \begin{cases}1, & m=1,\\ 0, & m=2,3,\cdots,L.\end{cases}$$

We can rewrite this equation in the following form

$$\begin{bmatrix} 1^0 & 3^0 & \cdots & (2L-1)^0 \\ 1^2 & 3^2 & \cdots & (2L-1)^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1^{2L-1} & 3^{2L-2} & \cdots & (2L-1)^{2L-2} \end{bmatrix}\begin{bmatrix} 1\beta_1 \\ 3\beta_2 \\ \vdots \\ (2L-1)\beta_L \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Now solving the above system, we get the following solutions

$$\beta_l = \frac{(-1)^{l+1}}{2l-1}\prod_{1\le m\le L, m\ne l}\left|\frac{(2m-1)^2}{(2l-1)^2-(2m-1)^2}\right|, \quad l=1,\cdots,L.$$

Journal of Applied Mathematics and Physics, 2019, 7, 1775-1785
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

The Aleksandrov Problem in Non-Archimedean 2-Fuzzy 2-Normed Spaces

Meimei Song, Haixia Jin*

Science of College, Tianjin University of Technology, Tianjin, China
Email: *173106301@stud.tjut.edu.cn

How to cite this paper: Song, M.M. and Jin, H.X. (2019) The Aleksandrov Problem in Non-Archimedean 2-Fuzzy 2-Normed Spaces. *Journal of Applied Mathematics and Physics*, **7**, 1775-1785.
https://doi.org/10.4236/jamp.2019.78121

Received: July 19, 2019
Accepted: August 16, 2019
Published: August 19, 2019

Abstract

We introduce the definition of non-Archimedean 2-fuzzy 2-normed spaces and the concept of isometry which is appropriate to represent the notion of area preserving mapping in the spaces above. And then we can get isometry when a mapping satisfies AOPP and (*) (in article) by applying the Benz's theorem about the Aleksandrov problem in non-Archimedean 2-fuzzy 2-normed spaces.

Keywords

Non-Archimedean 2-Fuzzy 2-Normed Space, Isometry, Benz's Theorem

1. Introduction

Let X, Y be two metric spaces. For a mapping $f: X \to Y$, for all $x_1, x_2 \in X$, if f satisfies,

$$d_Y\left(f\left(x_1\right), f\left(x_2\right)\right) = d_X\left(x_1, x_2\right)$$

where $d_X(\cdot,\cdot), d_Y(\cdot,\cdot)$ denote the metrics in the spaces X, Y, then f is called an isometry. It means that for some fixed number $p > 0$, assume that f preserves distance p; *i.e.*, for all x_1, x_2 in X, if $d_X\left(x_1, x_2\right) = p$, we can get $d_Y\left(f\left(x_1\right), f\left(x_2\right)\right) = p$. Then we say p is a conservative distance for the mapping f. Whether there exists a single conservative distance for some f such that f is an isometry from X to Y, is the basic issue of conservative distances. It is called the Aleksandrov problem.

Theorem 1.1. ([1]) *Let X, Y be two real normed linear spaces (or NLS) with $\dim X > 1$, $\dim Y > 1$ and Y is strictly convex, assume that a fixed real number $p > 0$ and that a fixed integer $N > 1$. Finally, if $f: X \to Y$ is a mapping satisfies*

1) $\|x_1 - x_2\| = p \Rightarrow \|f(x_1) - f(x_2)\| \le p$

2) $\|x_1 - x_2\| = N \cdot p \Rightarrow \|f(x_1) - f(x_2)\| \ge N \cdot p$

for all $x_1, x_2 \in X$. Then f is an affine isometry. we can call Benz's theorem.

We can see some results about the Aleksandrov problem in different spaces in [2]-[10]. A natural question is that: Whether the Aleksandrov problem can be proved in non-Archimedean 2-fuzzy 2-normed spaces under some conditions. So in this article, we will give the definition of non-Archimedean 2-fuzzy 2-normed spaces according to [11] [12] [13] [14], then by applying the Benz's theorem to fix the value of p and N to solve problems.

If a function from a field K to $[0,\infty)$ satisfies

(T_1) $|a| \ge 0, |a| = 0 \Leftrightarrow a = 0$;

(T_2) $|ab| = |a||b|$;

(T_3) $|a+b| \le \max\{|a|, |b|\}$.

for all $a, b \in K$, then the field K is called a non-Archimedean field.

We can know $|-1| = |1| = 1$, $|a| \le 1$ for all $a \in N$ from the above definition. An example of a non-Archimedean valuation (or NAV) is the function $|\cdot|$ taking $|0| = 0$ and others into 1.

In 1897, Hence in [15] found that p-adic numbers play a vital role in the complex analysis, the norm derived from p-adic numbers is the non-Archimedean norm, the analysis of the non-Archimedean has important applications in physics.

Definition 1.2. *Let X be a vector space and dim $X \ge 2$. A function $\|\cdot,\cdot\| : X \to [0,\infty)$ is called non-Archimedean 2-norm, if and only if it satisfies*

(T_1) $\|x_1, x_2\| \ge 0$, $\|x_1, x_2\| = 0$ iff x_1, x_2 are linearly dependent;

(T_2) $\|x_1, x_2\| = \|x_2, x_1\|$;

(T_3) $\|rx_1, x_2\| = |r|\|x_1, x_2\|$;

(T_4) $\|x_1 + x_2, y\| \le \max\{\|x_1, y\|, \|x_2, y\|\}$

for all $x_1, x_2, y \in X, r \in K$. Then $(X, \|\cdot,\cdot\|)$ is called non-Archimedean 2-normed space over the field K.

Definition 1.3. *An NAV $|\cdot|$ in a linear space X over a field K. A function $F : X \times \mathbb{R} \to [0,1]$ is said to be a non-Archimedean fuzzy norm on X, if and only if for all $x, x_1, x_2 \in X$ and $s, t \in \mathbb{R}$,*

(F1) $F(x,s) = 0$ with $s \le 0$,

(F2) $F(x,s) = 1$ iff $x = 0$ for all $s > 0$,

(F3) $F(cx,s) = F\left(x, \dfrac{s}{|c|}\right)$, for $c \ne 0$ and $c \in K$,

(F4) $F(x_1 + x_2, s+t) \ge \min\{F(x_1,s), F(x_2,t)\}$,

(F5) $F(x,*)$ is a nondecreasing function of $s \in R$ and $\lim_{s\to\infty} F(x,s) = 1$.

Then (X,F) is known as a non-Archimedean fuzzy normed space (or F-NANS).

Theorem 1.4. *Let (X,F) be an F-NANS. Assume the condition that*:

(F6) $F(x,s) > 0$ for all $s > 0 \Rightarrow x = 0$.

Define $\|x\|_\alpha = \inf\{s : F(x,s) \ge \alpha\}, \alpha \in (0,1)$. We call these α-norms on X or the fuzzy norm on X.

Proof: 1) Let $\|x\|_\alpha = 0$, it implies that $\inf\{s : F(x,s) \geq \alpha\} = 0$, then for all $s \in R$, $s > 0$, $F(x,s) \geq \alpha > 0$, so $x = 0$;

Conversely, assume that $x = 0$, by (F2), $F(x,s) = 1$ for all $s > 0$, then $\inf\{s : F(x,s) \geq \alpha\} = 0$ for all $\alpha \in (0,1)$, so $\|x\|_\alpha = 0$.

2) By (F3), if $c \neq 0$, then

$$\|cx\|_\alpha = \inf\{s : F(cx,s) \geq \alpha\} = \inf\left\{s : F\left(x, \frac{s}{|c|}\right) \geq \alpha\right\}$$

Let $t = \frac{s}{|c|}$, then

$$\|cx\|_\alpha = \inf\{|c|t : F(x,t) \geq \alpha\} = |c|\inf\{t : F(x,t) \geq \alpha\} = |c| \cdot \|x\|_\alpha$$

If $c = 0$, then

$$\|cx\|_\alpha = 0 = c\|x\|_\alpha$$

3) We have

$$\begin{aligned}
&\max\{\|x\|_\alpha, \|y\|_\alpha\} \\
&= \max\{\inf\{s, F(x,s) \geq \alpha\}, \inf\{t, F(x,t) \geq \alpha\}\} \\
&= \inf\{\max\{s,t\}, F(x,s) \geq \alpha, F(x,t) \geq \alpha\} \\
&\geq \inf\{s+t, F(x+y, s+t) \geq F(x+y, \max\{s,t\})\} \\
&\quad \geq \min\{F(x,s) \geq \alpha, F(x,t) \geq \alpha\} \geq \alpha\} \\
&\geq \inf\{r, F(x+y,r) \geq \alpha\} = \|x+y\|_\alpha
\end{aligned}$$

□

Example 1.5. *Let $(X, \|\cdot\|)$ be a non-Archimedean normed space. Define*

$$F(x,s) = \begin{cases} \dfrac{s}{s+\|x\|}, & s > 0, \\ 0, & s \leq 0. \end{cases}$$

for all $x \in X$, Then (X,F) is a F-NANS.

Definition 1.6. *Let Z be any non-empty set and $\Im(Z)$ be the set of all fuzzy sets on Z. For $Z_1, Z_2 \in \Im(Z)$ and $\lambda \in K$, define*

$$Z_1 + Z_2 = \{(z_1 + z_2, \mu_1 \wedge \mu_2) \mid (z_1, \mu_1) \in Z_1, (z_2, \mu_2) \in Z_2\}$$

and

$$\lambda Z_1 = \{(\lambda z_1, \mu_1) \mid (z_1, \mu_1) \in Z_1\}$$

Definition 1.7. *A non-Archimedean fuzzy linear space $\hat{X} = X \times (0,1]$ over the field K, we define the addition and scalar multiplication operation of X as following:* $(x_1, \mu_1) + (x_2, \mu_2) = (x_1 + x_2, \mu_1 \wedge \mu_2)$, $\lambda(x_1, \mu_1) = (\lambda x_1, \mu_1)$, *if for every $(x_1, \mu_1) \in X$, we have a related non-negative real numebr, $\|(x_1, \mu_1)\|$ is the fuzzy norm of (x_1, μ_1) in such that*

(T_1) $\|(x_1, \mu_1)\| = 0 \Leftrightarrow x_1 = 0, \mu_1 \in (0,1]$;

(T2) $\|\lambda(x_1,\mu_1)\| = |\lambda|\|(x_1,\mu_1)\|$;

(T3) $\|(x_1,\mu_1)+(x_2,\mu_2)\| \le \max\{\|(x_1,\mu_1\wedge\mu_2)\|,\|(x_2,\mu_1\wedge\mu_2)\|\}$;

(T4) $\|(x_1,\vee_t\mu_t)\| = \wedge_t\|(x_1,\mu_t)\|$ for all $\mu_t\in(0,1]$.

for every $(x_1,\mu_1),(x_2,\mu_2)\in X, \lambda\in K$, then we say that X is an F-NANS.

Definition 1.8. *Let X be a non-empty non-Archimedean field set, $\Im(X)$ be the set of all fuzzy sets on X. If $f_1\in\Im(X)$, then $f_1=\{(x_1,\mu_1): x_1\in X, \mu_1\in(0,1]\}$. Clearly, $|f_1(x_1)|\le 1$, so f_1 is a bounded function. Let $K\in\mathbb{Q}$, then $\Im(X)$ is a non-Archimedean linear space over the field K and the addition, scalar multiplication are defined as follows*

$$f_1+f_2=\{(x_1,\mu_1)+(x_2,\mu_2)\}=\{(x_1+x_2,\mu_1\wedge\mu_2)\,|\,(x_1,\mu_1)\in f_1,(x_2,\mu_2)\in f_2\}$$

and

$$\lambda f_1=\{(\lambda x_1,\mu_1)\,|\,(x_1,\mu_1)\in f_1\}$$

If for every $f\in\Im(X)$, there is a related non-negative real number $\|f\|$ called the norm of f in such that for all $f_1=(x_1,\mu_1), f_2=(x_2,\mu_2)\in\Im(X)$

(T1) $\|f\|=0$ iff $f=0$. For

$$\|f\|=\{\|(x_1,\mu_1)\|\}=0$$
$$\Leftrightarrow x_1=0,\mu_1\in(0,1]$$
$$\Leftrightarrow f=0.$$

(T2) $\|\lambda f\|=|\lambda|\|f\|, \lambda\in K$. For

$$\|\lambda f\|=\{\|\lambda(x_1,\mu_1)\|\}=\{|\lambda|\|(x_1,\mu_1)\|\}=|\lambda|\|f\|$$

(T3) $\|f_1+f_2\|\le\max\{\|f_1\|,\|f_2\|\}$. For

$$\begin{aligned}\|f_1+f_2\| &= \{\|(x_1,\mu_1)+(x_2,\mu_2)\|\}\\ &= \{\|(x_1+x_2),(\mu_1\wedge\mu_2)\|\}\\ &\le \max\{\|(x_1,\mu_1\wedge\mu_2)\|,\|(x_2,\mu_1\wedge\mu_2)\|\}\\ &\le \max\{\|f_1\|,\|f_2\|\}\end{aligned}$$

Then the linear space $\Im(X)$ is a non-Archimedean normed space.

Definition 1.9. ([4]) *A 2-fuzzy set on X is a fuzzy set on $\Im(X)$.*

Definition 1.10. *A NAV $|\cdot,\cdot|$ in a linear space $\Im(X)$ over a field K. If a function $F:\Im(X)^2\times\mathbb{R}\to[0,1]$ is a non-Archimedean 2-fuzzy 2-norm on X (or a fuzzy 2-norm on $\Im(X)$), iff for all $f_1,f_2,f_3\in\Im(X)$, $s,t\in\mathbb{R}$,*

(F1) $F(f_1,f_2,s)=0$ for $s\le 0$;

(F2) $F(f_1,f_2,s)=1$ iff f_1,f_2 are linearly dependent for all $s>0$;

(F3) $F(f_1,f_2,s)=N(f_2,f_1,s)$;

(F4) $F(cf_1,f_2,s)=N\left(f_1,f_2,\dfrac{s}{|c|}\right)$, for $c\ne 0$ and $c\in K$;

(F5) $F(f_1,f_2+f_3,s+t)\ge\min\{F(f_1,f_2,s),F(f_1,f_3,t)\}$;

(F6) $F(f_1, f_2, *)$ is a nondecreasing function of R and $\lim_{s\to\infty} F(f_1, f_2, s) = 1$;

Then $(\Im(X), F)$ is called a non-Archimedean fuzzy 2-normed space (or FNA-2) or (X, F) is a non-Archimedean 2-fuzzy 2-normed space.

Theorem 1.11. *Let* $(\Im(X), F)$ *be an FNA-2. Suppose the condition that*:

(F7) $N(f_1, f_2, s) > 0$ for all $s > 0 \Rightarrow f_1$ and f_2 are linearly dependent.

Define $\|f_1, f_2\|_\alpha = \inf\{t : N(f_1, f_2, s) \geq \alpha, \alpha \in (0,1)\}$. We call these α-2-norms on $\Im(X)$ or the 2-fuzzy 2-norm on X.

Proof: It is similar to the proof of Theorem 1.4. □

2. Main Result

From now on, if we have no other explanation, let $\dim \Im(X) \geq 2$, $\dim \Im(Y) \geq 2$. $\blacktriangle = \|f-h, g-h\|_\alpha$, $\blacktriangledown = \|\psi(f)-\psi(h), \psi(g)-\psi(h)\|_\beta$

Definition 2.1. *Let* $\Im(X), \Im(Y)$ *be two FNA-2 and a mapping* $\psi : \Im(X) \to \Im(Y)$. *If for all* $f, g, h \in \Im(X)$ *and* $\alpha, \beta \in (0,1)$, *we have*

$$\|\psi(f)-\psi(h), \psi(g)-\psi(h)\|_\beta = \|f-h, g-h\|_\alpha \quad (\nabla)$$

then ψ is called 2-isometry.

Definition 2.2. *For a mapping* $\psi : \Im(X) \to \Im(Y)$ *and* $f, g, h \in \Im(X)$

1) If $\blacktriangle = 1$, then $\blacktriangledown = 1$, we say ψ satisfies the area one preserving property (AOPP).

2) If $\blacktriangle = n$, then $\blacktriangledown = n$, we say ψ satisfies the area *n* for each *n* (AnPP).

Definition 2.3. *We say a mapping* $\psi : \Im(X) \to \Im(Y)$ *preserves collinear, if* f, g, h *mutually disjoint elements of* $\Im(X)$, *then exist some real number t we have*

$$\psi(g)-\psi(h) = t(\psi(f)-\psi(h))$$

Next, we denote $\|\psi(f)-\psi(h), \psi(g)-\psi(h)\|_\beta \leq \|f-h, g-h\|_\alpha \quad (*)$.

Lemma 2.4. *Let* $\Im(X)$ *and* $\Im(Y)$ *be two FNA-2. If* $\blacktriangle \leq 1$, *a mapping* $\psi : \Im(X) \to \Im(Y)$ *satisfies* $(*)$ *and AOPP, then we can get* (∇) *where* $\blacktriangle \leq 1$.

Proof: 1) Firstly, we prove that *f* preserves collinear. We assume that $\blacktriangle = 0$, according to $(*)$, we get

$$\|\psi(f)-\psi(h), \psi(g)-\psi(h)\|_\beta = 0$$

then $\psi(f)-\psi(h)$ and $\psi(g)-\psi(h)$ are linearly dependent. So we obtain that ψ preserves collinear.

2) Secondly, we prove that when $\blacktriangle \leq 1$, we can get (∇).

If

$$\blacktriangledown < \blacktriangle$$

Let $\omega = h + \dfrac{f-h}{\|f-h, g-h\|_\alpha}$, then $\|\omega-h, g-h\|_\alpha = 1$, so

$$\|\psi(\omega)-\psi(h),\psi(g)-\psi(h)\|_\beta = 1 \quad (\Delta)$$

Since

$$\|\omega - f, g-h\|_\alpha = \left\| \frac{f-h}{\|f-h, g-h\|} - (f-h), g-h \right\| = 1-\blacktriangle$$

according to $(*)$, we have

$$\|\psi(\omega)-\psi(f),\psi(g)-\psi(h)\|_\beta \le \|\omega - f, g-h\|_\alpha = 1-\blacktriangle$$

Since f preserves collinear, so there exists a real number s such that

$$\psi(\omega)-\psi(h) = s(\psi(f)-\psi(h))$$

and

$$\psi(\omega)-\psi(f) = (s-1)(\psi(f)-\psi(h))$$

So, we get

$$\begin{aligned}
&\|\psi(\omega)-\psi(h),\psi(g)-\psi(h)\|_\beta \\
&= |s|\blacktriangledown \\
&\le |s-1|\blacktriangledown + \blacktriangledown \\
&= \|\psi(\omega)-\psi(f),\psi(g)-\psi(h)\|_\beta + \blacktriangledown \\
&< 1-\blacktriangle+\blacktriangle = 1
\end{aligned}$$

This contradicts with Δ. □

Lemma 2.5. *Let* $\Im(X)$ *and* $\Im(Y)$ *be two FNA-2. If a mapping* $\psi: \Im(X) \to \Im(Y)$ *satisfies AOPP and preserves collinear, then*

1) ψ is an injective;

2) if $\phi(f) = \psi(f) - \psi(0)$, then $\phi(f+g) = \phi(f) + \phi(g)$ and $\phi(\lambda f) = \lambda\phi(f)$ with $0 < \lambda < 1$.

Proof: 1) We prove ψ is injective. Let $f, g \in \Im(X)$, since dim $\Im(X) \ge 2$, there exists an element $h \in \Im(X)$ such that $f-h, g-h$ are linearly independent. Hence $\blacktriangle \ne 0$.

Let $\gamma = h + \dfrac{g-h}{\|f-h, g-h\|_\alpha}$, then $\|f-h, \gamma-h\|_\alpha = 1$, and ψ satisfies AOPP, so

$$\|\psi(f)-\psi(h),\psi(\gamma)-\psi(h)\|_\beta = 1$$

we can see $\psi(h) \ne \psi(f)$. So the mapping ψ is injective.

2) Let f, g, h mutually disjoint elements of $\Im(X)$ and $f = \dfrac{g+h}{2}$, so $f-h = g-f$ $(\star)$. Since ψ is injective and preserves collinear, there exist $s \ne 0$ such that

$$\psi(g)-\psi(f) = s(\psi(h)-\psi(f))$$

Since dim $\Im(X) \ge 2$, there exist an element $f_1 \in \Im(X)$ such that

$\|g-f, f_1-f\|_\alpha \neq 0$. Let $\eta = f + \dfrac{f_1 - f}{\|g-f, f_1-f\|_\alpha}$, then $\|g-f, \eta-f\|_\alpha = 1$ and

$$\|\psi(g)-\psi(f), \psi(\eta)-\psi(f)\|_\beta = 1.$$

So,

$$\|\psi(h)-\psi(f), \psi(\eta)-\psi(f)\|_\beta = \left|\frac{1}{s}\right|.$$

Since $(\star)$, we get $\|h-f, \eta-f\|_\alpha = 1$ and

$$\|\psi(h)-\psi(f), \psi(\eta)-\psi(f)\|_\beta = 1.$$

According to the mapping ψ is injective, so $s = -1$, and

$$\psi\left(\frac{g+h}{2}\right) = \frac{\psi(g)+\psi(h)}{2}$$

Let $\phi(f) = \psi(f) - \psi(0)$, so we have

$$\phi\left(\frac{g+h}{2}\right) = \frac{\phi(g)+\phi(h)}{2}$$

Therefore

$$\phi\left(\frac{f}{2}\right) = \phi\left(\frac{f+0}{2}\right) = \frac{\phi(f)}{2}$$

and

$$\phi(f+g) = \phi\left(\frac{2f+2g}{2}\right) = \frac{\phi(2f)}{2} + \frac{\phi(2g)}{2} = \phi(f) + \phi(g)$$

So ϕ is additive.

From the lemma 2.4, we know that if $\blacktriangle \leq 1$, then ϕ satisfies 2-isometry.

$$0 = \|\lambda f, f\|_\alpha = \|\psi(\lambda f)-\psi(0), \psi(f)-\psi(0)\|_\beta = \|\phi(\lambda f), \phi(f)\|_\beta$$

so $\phi(\lambda f)$ and $\phi(f)$ is linearly dependent *i.e.* $\phi(\lambda f) = s\phi(f)$.

Next we assume $\|f, g\|_\alpha = \lambda$,

$$\frac{1}{\lambda}\|f, g\|_\alpha = \left\|\frac{f}{\lambda} - 0, g - 0\right\|_\alpha = 1$$

and

$$\begin{aligned} 1 &= \left\|\phi\left(\frac{f}{\lambda}\right) - \phi(0), \phi(g) - \phi(0)\right\|_\beta \\ &= \left\|\phi\left(\frac{f}{\lambda}\right), \phi(g)\right\|_\beta \\ &= \frac{1}{|s|}\|\phi(f), \phi(g)\|_\beta \\ &= \frac{1}{|s|}\|f, g\|_\alpha \end{aligned}$$

Thus $\lambda = |s|$, if $s = -\lambda$, then $\phi(\lambda f) = -\lambda\phi(f)$, but

$$\begin{aligned}|\lambda-1|\|f,g\|_\alpha &= \|\lambda f-f,g-0\|_\alpha \\ &= \|\phi(\lambda f)-\phi(f),\phi(g)-\phi(0)\|_\beta \\ &= \|-\lambda\phi(f)-\phi(f),\phi(g)-\phi(0)\|_\beta \\ &= (\lambda+1)\|\phi(f),\phi(g)\|_\beta \\ &= (\lambda+1)\|f,g\|_\alpha\end{aligned}$$

so $|\lambda-1|=(\lambda+1)$. It contradicts with $0<\lambda<1$. Thus $\phi(\lambda f)=\lambda\phi(f)$. □

Lemma 2.6. *Let* $\Im(X)$ *and* $\Im(Y)$ *be FNA-2. If* $\blacktriangle\le 1$, *a mapping* $\psi:\Im(X)\to\Im(Y)$ *satisfies* $(*)$ *and AOPP, then we can get for all* $f,g,h\in\Im(X)$, *we can get* (∇).

Proof: From lemma 2.4, we know ψ preserves collinear.

For any $f,g,h\in\Im(X)$, there exist two numbers $m,n\in\mathbb{N}^*$ such that $\blacktriangle\le\frac{m}{n}$.

So,

$$\left\|\psi\left(\frac{f}{m}\right)-\psi\left(\frac{h}{m}\right),\psi(g)-\psi(h)\right\|_\beta \le \left\|\frac{f-h}{m},g-h\right\|_\alpha \le \frac{1}{n}$$

and

$$\left\|\left(\frac{f}{m}\right)-\phi\left(\frac{h}{m}\right),\phi(g)-\phi(h)\right\|_\beta \le \frac{1}{n}$$

By lemma 2.5, we have

$$\left\|\frac{1}{m}(\phi(f)-\phi(h)),\phi(g)-\phi(h)\right\|_\beta \le \frac{1}{n}$$

$$\|\phi(f)-\phi(h),\phi(g)-\phi(h)\|_\beta \le \frac{m}{n}$$

Thus

$$\|\psi(f)-\psi(h),\psi(g)-\psi(h)\|_\beta \le \frac{m}{n}$$

□

Lemma 2.7. *Let* $\Im(X)$ *and* $\Im(Y)$ *be two FNA-2. If a mapping* $\psi:\Im(X)\to\Im(Y)$ *satisfies AOPP and* $(*)$ *for all* $f,g,h\in\Im(X)$ *with* $\blacktriangle\le 1$, *then* ψ *satisfies AnPP.*

Proof: Let $f,g,h\in\Im(X)$ and $n\in\mathbf{N}$. Let

$$\blacktriangle=n,\ g_i=h+\frac{i}{n}(g-h)$$

and

$$\|f-h,g_{i+1}-g_i\|_\alpha=1,\ i=0,1,\cdots,n-1.$$

So,

$$\|\psi(f)-\psi(h),\psi(g_{i+1})-\psi(g_i)\|_\beta=1,\ i=0,1,\cdots,n-1.$$

We know ψ preserves collinear. So there exist a number $t \in \mathbb{R}$ such that

$$\psi(g_2)-\psi(g_1)=t\left(\psi(g_1)-\psi(g_0)\right)$$

Therefore

Then we have $t=\pm 1$. By lemma 2.5, $t=1$, so

$$\psi(g_2)-\psi(g_1)=\psi(g_1)-\psi(g_0).$$

In the same way, we can get

$$\psi(g_{i+1})-\psi(g_i)=\psi(g_i)-\psi(g_{i-1}),\ i=0,1,\cdots,n-1.$$

Hence

$$\begin{aligned}\psi(g)-\psi(h)&=\psi(g_n)-\psi(g_0)\\&=\psi(g_n)-\psi(g_{n-1})+\psi(g_{n-1})-\psi(g_{n-2})+\cdots+\psi(g_1)-\psi(g_0)\\&=n\left(\psi(g_1)-\psi(g_0)\right)\end{aligned}$$

Therefore

$$\begin{aligned}\blacktriangledown&=\left\|\psi(f)-\psi(h),n\left(\psi(g_1)-\psi(g_0)\right)\right\|_\beta\\&=n\left\|\psi(f)-\psi(h),\psi(g_1)-\psi(g_0)\right\|_\beta=n\end{aligned}$$

□

Theorem 2.8. *Let* $\Im(X)$ *and* $\Im(Y)$ *be two FNA-2. If a mapping* $\psi:\Im(X)\to\Im(Y)$ *satisfies AOPP and* $(*)$ *for all* $f,g,h\in\Im(X)$ *with* $\blacktriangle\le 1$, *then* ψ *is* 2*-isometry.*

Proof: Since lemma 2.4, we just need to prove that (∇) with $\blacktriangle>1$.

We can assume that when $\blacktriangle>1$, for all $f,g,h\in\Im(X)$, we have $\blacktriangledown<n_0+1$. and there exist a number $n_0\in\mathbb{N}^*$ such that

Let $\tau=f+\dfrac{n_0+1}{\|f-h,g-h\|_\alpha}(f-h)$, then

$$\|\tau-f,g-h\|_\alpha=n_0+1$$

and

$$\|\tau-h,g-h\|_\alpha=n_0+1-\blacktriangle$$

Since ψ preserves collinear, there exist a number $c\in\mathbb{R}$ such that

$$\psi(\tau)-\psi(f)=c\left(\psi(h)-\psi(f)\right)$$

Since 2),

$$\begin{aligned}n_0+1&=\left\|\psi(\tau)-\psi(f),\psi(g)-\psi(h)\right\|_\beta\\&=|c|\blacktriangledown\\&\le|c-1|\blacktriangledown+\blacktriangledown\\&=\left\|\psi(\tau)-\psi(h),\psi(g)-\psi(h)\right\|_\beta+\blacktriangledown\\&<n_0+1-\blacktriangle+\blacktriangle=n_0+1\end{aligned}$$

which is contradiction, so

$$\blacktriangledown \geq n_0 + 1$$

Therefore, we get (∇) with $\blacktriangle > 1$. Hence

$$\left\|\psi(f)-\psi(h),\psi(g)-\psi(h)\right\|_\beta = \left\|f-h,g-h\right\|_\alpha$$

for all $f,g,h \in \Im(X)$. □

Conflicts of Interest

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

References

[1] Benz, W. and Berens, H. (1987) A Contribution to a Theorem of Ulam and Mazur. *Aequationes Mathematicae*, **34**, 61-63. https://doi.org/10.1007/BF01840123

[2] Bag, T. and Samanta, S.K. (2003) Finite Dimensional Fuzzy Normed Linear Spaces. *The Journal of Fuzzy Mathematics*, **11**, 687-705.

[3] Chu, H.Y., Park, C.G. and Park, W.G. (2004) The Aleksandrow Problem in Linear 2-Normed Spaces. *Journal of Mathematical Analysis and Applications*, **289**, 666-672. https://doi.org/10.1016/j.jmaa.2003.09.009

[4] Somasundaram, R.M. and Beaula, T. (2009) Some Aspects of 2-Fuzzy 2-Normed Linear Spaces. *Bulletin of the Malaysian Mathematical Sciences Society*, **32**, 211-221.

[5] Zheng, F.H. and Ren, W.Y. (2014) The Aleksandrow Problem in Quasi Convex Normed Linear Space. *Acta Scientiarum Natueralium University Nankaiensis*, No. 3, 49-56.

[6] Huang, X.J. and Tan, D.N. (2017) Mapping of Conservative Distances in p-Normed Spaces ($0 < p \leq 1$). *Bulletin of the Australian Mathematical Society*, **95**, 291-298. https://doi.org/10.1017/S0004972716000927

[7] Ma, Y.M. (2000) The Aleksandrov Problem for Unit Distance Preserving Mapping. *Acta Mathematica Science*, **20B**, 359-364. https://doi.org/10.1016/S0252-9602(17)30642-2

[8] Wang, D.P., Liu, Y.B. and Song, M.M. (2012) The Aleksandrov Problem on Non-Archimedean Normed Spaces. *Arab Journal of Mathematical Science*, **18**, 135-140. https://doi.org/10.1016/j.ajmsc.2011.10.002

[9] Ma, Y.M. (2016) The Aleksandrov-Benz-Rassias Problem on Linear n-Normed Spaces. *Monatshefte für Mathematik*, 180, 305-316. https://doi.org/10.1007/s00605-015-0786-8

[10] Huang, X.J. and Tan, D.N. (2018) Mappings of Preserves n-Distance One in N-Normed Spaces. *Aequations Mathematica*, **92**, 401-413. https://doi.org/10.1007/s00010-018-0539-6

[11] Xu, T.Z. (2013) On the Mazur-Ulanm Theorem in Non-Archimedean Fuzzy n-Normed Spaces. *ISRN Mathematical Analysis*, **67**, 1-7. https://doi.org/10.1155/2013/814067

[12] Chang, L.F. and Song, M.M. (2014) On the Mazur-Ulam Theorem in Non-Archimedean Fuzzy 2-Normed Spaces. *Mathematica Applicata*, **27**, 355-359.

[13] Alaca (2010) New Perspective to the Mazur-Ulam Problem in 2-Fuzzy 2-Normed Linear Spaces. *Iranian Journal of Fuzzy Systems*, **7**, 109-119.

[14] Park, C. and Alaca, C. (2013) Mazur-Ulam Theorem under Weaker Conditions in the Framework of 2-Fuzzy 2-Normed Linear Spaces. *Journal of Inequalities and Applications*, **2018**, 78. https://doi.org/10.1186/1029-242X-2013-78

[15] Hensel, K. (1897) Über eine neue Begründung der Theorie der algebraischen Zahlen. *Jahresbericht der Deutschen Mathematiker- Vereinigung*, **6**, 83-88.

Journal of Applied Mathematics and Physics, 2019, 7, 1786-1799
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Chi-Square Distribution: New Derivations and Environmental Application

Thomas M. Semkow[1,2], Nicole Freeman[3], Umme-Farzana Syed[1], Douglas K. Haines[1], Abdul Bari[1], Abdul J. Khan[1], Kimi Nishikawa[1], Adil Khan[1], Adam G. Burn[1,2], Xin Li[1,2], Liang T. Chu[1,2]

[1]Wadsworth Center, New York State Department of Health, Albany, NY, USA
[2]Department of Environmental Health Sciences, University at Albany, State University of New York, Rensselaer, NY, USA
[3]Averill Park Central School District, Averill Park, NY, USA
Email: thomas.semkow@health.ny.gov

How to cite this paper: Semkow, T.M., Freeman, N., Syed, U.-F., Haines, D.K., Bari, A., Khan, A.J., Nishikawa, K., Khan, A., Burn, A.G., Li, X. and Chu, L.T. (2019) Chi-Square Distribution: New Derivations and Environmental Application. *Journal of Applied Mathematics and Physics*, **7**, 1786-1799.
https://doi.org/10.4236/jamp.2019.78122

Received: July 19, 2019
Accepted: August 16, 2019
Published: August 19, 2019

Abstract

We describe two new derivations of the chi-square distribution. The first derivation uses the induction method, which requires only a single integral to calculate. The second derivation uses the Laplace transform and requires minimum assumptions. The new derivations are compared with the established derivations, such as by convolution, moment generating function, and Bayesian inference. The chi-square testing has seen many applications to physics and other fields. We describe a unique version of the chi-square test where both the variance and location are tested, which is then applied to environmental data. The chi-square test is used to make a judgment whether a laboratory method is capable of detection of gross alpha and beta radioactivity in drinking water for regulatory monitoring to protect health of population. A case of a failure of the chi-square test and its amelioration are described. The chi-square test is compared to and supplemented by the *t*-test.

Keywords

Mathematical Induction, Laplace Transform, Gamma Distribution, Chi-Square Test, Gross Alpha-Beta, Drinking Water

1. Introduction

The chi-square distribution (CSD) has been one of the most frequently used distributions in science. It is a special case of the gamma distribution (see Section 2). The latter has been an important distribution in fundamental physics, for example as kinetic energy distribution of particles in an ideal gas (Maxwell-Boltzmann) [1] or the kinetic energy distribution of particles emitted from excited nuclei in nuclear reactions [2]. A historical context for the development of the CSD is de-

scribed in References [3] and [4]. Its first derivation is attributed to Bienaymé [5], who used multiple integrals over normal variables and substitutions. Abbe [6] used a method of integration in the complex plane to solve multiple integrals. The most general derivation is attributed to Helmert, who proposed a classic transformation to derive CSD, including calculation of the Jacobian determinant of transformation [7]. This transformation can be worked out into polar variables, which is described in statistical textbooks [4] [8].

The established fundamental derivations of the CSD described above lend themselves to complicated handling of multiple integrals. On the contrary, the simplified derivations use the fact that CSD is a special case of the gamma distribution. Owing to the integrable and recursive properties of the gamma distribution, as well as its moment generating function (Mgf), simplified derivations of CSD are described in the textbooks [9] [10]. Another simplified derivation uses Bayesian inference [11]. In Section 2, we refer to these methods for comparisons.

In this work, we present two new methods of derivation of the CSD. They are both within the simplified category. One of them is mathematical induction. The original derivation was done by Helmert [12] using a 2-step forward mathematical induction. We have elaborated on that and observed that the CSD has certain recursive property, which enables its derivation using a single-step induction plus the well-known theorem for beta and gamma functions. Another derivation method we describe is by the Laplace transform. This method has some similarity to the Mgf and characteristic function methods, owing to the presence of exponentiation. It uses a complex-variable integration and it is free from many assumptions of the other methods. The two new derivations of the CSD by mathematical induction and Laplace transform are described in Section 2.

Chi-square testing (CST) is closely related to and based upon the CSD. It has its origins in the discovery of the goodness-of-fit test by Pearson [13]. In the goodness-of-fit, one calculates the test statistics as

$$\chi_\nu^2 = \sum_{i=1}^{m} \frac{(O_i - E_i)^2}{E_i}, \tag{1}$$

where O_i is frequency of observation, E_i is expected frequency based on an assumed model distribution, for category of type i, and m is the number of categories. Both O_i and E_i are unitless. $\nu = m - 1 - p$ is the number of degrees of freedom, where p is number of parameters of the model distribution calculated from the data. For any model distribution, Equation (1) leads asymptotically to the CSD when the number of observations is large, which has been proved for the multinomial distribution by Pearson [13]. The goodness-of-fit CST has been extensively used in statistics and widely applied to many fields [3] [14]. It is worth noting that the interpretation of the degrees of freedom was provided by Fisher [15]. As example in physics, CST goodness-of-fit has been used to verify Poisson fluctuations of radioactivity counter [14] [16].

Another form of the chi-square variable from Equation (1) is written in the

general form as

$$\chi_\nu^2 = \sum_{i=1}^{n} \left(\frac{x_i - \mu_i}{\sigma_i} \right)^2, \tag{2}$$

where n is the number of observations, x_i is the observed variable, μ_i is the expected value, σ_i is the standard deviation, and $\nu \le n$. The variables in Equation (2) can be expressed in physical units. In the limit of large number of observations, the variable and parameters of Equation (2) are approximated by those of the normal variates, and the χ_ν^2 distributes as CSD. In this work, we generalize this CST test to a combined test for variance and location as well as verify it with the *t*-test [17]. The test statistics studied are described in Section 3.

Within the context of this work, we present a unique application of the CST to the detection of radioactive contaminants in drinking water required by the Safe Drinking Water Act (SDWA) in the US. The bulk of natural alpha and beta/gamma (photon) radioactivity in drinking water originates from the possible presence of ^{238}U and ^{232}Th natural radioactive-series progeny, ^{226,228}Ra and their progeny, as well as ^{40}K radionuclides [18]. The SDWA regulations [19] establish a Maximum Contaminant Level (MCL) of 15 pCi/L (555 mBq/L) for gross alpha (GA) radioactivity, excluding U and Rn. For gross beta (GB) radioactivity, the MCL is limited by the total body or any organ radiation dose of 4 mrem/y (40 μSv/y). For both GA and GB, the Maximum Contaminant Level Goal (MCLG) is zero. Furthermore, SDWA requires Detection Limits (DL) of 3 pCi/L (111 mBq/L) and 4 pCi/L (148 mBq/L) for GA and GB radioactivity, respectively. These DLs must be met by all public health laboratories accredited for monitoring of GA and GB radioactivity in drinking water in the US. In Section 4, we detail a CST procedure to verify if the required above-mentioned DLs are met [20]. We investigate the reasons and consequences of failed CST and ameliorate such cases.

2. Chi-Square Distribution

The probability density function (Pdf) of the CSD is given by

$$\mathrm{Pdf}\left(\chi_\nu^2 \mid \nu\right) = \frac{\left(\chi_\nu^2\right)^{\nu/2-1} \mathrm{e}^{-\chi_\nu^2/2}}{2^{\nu/2}\Gamma\left(\nu/2\right)}, \tag{3}$$

where Γ is the gamma function. The expectation value of CSD is $E\left[\chi^2\right] = \nu$, and the variance $\mathrm{Var}\left[\chi^2\right] = 2\nu$ [21]. The CSD is a special case of the gamma distribution abbreviated as $\mathrm{gamma}\left(\chi_\nu^2 \mid a, b\right)$ with the parameters $a = \nu/2$ and $b = 2$ [21].

To derive Equation (3), we start with the general definition of χ_ν^2 statistics given by Equation (2) assuming normal variates. For a single normal variable x_1 with $\mathrm{Pdf}\left(x_1\right)$, the probability of $x_1 \in \left[x_1, x_1 + \mathrm{d}x_1\right]$ is given by

$$\mathrm{Pdf}\left(x_1\right)\mathrm{d}x_1 = \frac{1}{\sqrt{2\pi}\sigma_1} \mathrm{e}^{-\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 \Big/ 2} \mathrm{d}x_1. \tag{4}$$

By substituting $\chi_1^2 = \left((x_1 - \mu_1)/\sigma_1\right)^2$, we obtain from Equation (4)

$$\begin{aligned}\text{Pdf}\left(\chi_1^2 \mid 1\right)\text{d}\chi_1^2 &= \frac{2}{\sqrt{2\pi}\sigma_1}\text{e}^{-\chi_1^2/2}\left|\frac{\text{d}x_1}{\text{d}\chi_1^2}\right|\text{d}\chi_1^2 = \frac{\left(\chi_1^2\right)^{1/2-1}\text{e}^{-\chi_1^2/2}}{2^{1/2}\Gamma\left(1/2\right)}\text{d}\chi_1^2 \\ &= \text{gamma}\left(\chi_1^2 \mid 1/2, 2\right)\text{d}\chi_1^2,\end{aligned} \tag{5}$$

which has the Pdf given by Equation (3) for $\nu = 1$. In deriving Equation (5), we also used $\Gamma\left(1/2\right) = \sqrt{\pi}$, whereas factor of 2 originated from the fact that the x_1 variable ranging from minus infinity to plus infinity has been substituted with the χ_1^2 variable ranging from zero to plus infinity.

Let us assume that the $n+1$ term with the normal x_{n+1} variable was added to Equation (2), and that this addition raised the number of degrees of freedom to $\nu+1$. Then,

$$\chi_{\nu+1}^2 = \chi_\nu^2 + \left(\frac{x_{n+1} - \mu_{n+1}}{\sigma_{n+1}}\right)^2. \tag{6}$$

Using the calculus for probability density functions [21],

$$\text{Pdf}\left(\chi_{\nu+1}^2 \mid \nu+1\right)\text{d}\chi_{\nu+1}^2 = \int_{-\infty}^{+\infty}\text{Pdf}\left(\chi_\nu^2 \mid \nu\right)\text{d}\chi_\nu^2\text{Pdf}\left(x_{n+1}\right)\text{d}x_{n+1}. \tag{7}$$

Let us define a new variable *z*, such as

$$\left(\frac{x_{n+1} - \mu_{n+1}}{\sigma_{n+1}}\right)^2 = \chi_{\nu+1}^2\left(1-z\right). \tag{8}$$

By realizing that $\text{d}\chi_{\nu+1}^2 = \text{d}\chi_\nu^2$, and performing all substitutions, the right side of Equation (7) can be rewritten as

$$\begin{aligned}&2\int_0^1\text{Pdf}\left(\chi_\nu^2 \mid \nu\right)\text{Pdf}\left(x_{n+1}\right)\left|\frac{\text{d}x_{n+1}}{\text{d}z}\right|\text{d}z\text{d}\chi_{\nu+1}^2 \\ &= \frac{\left(\chi_{\nu+1}^2\right)^{(\nu+1)/2-1}\text{e}^{-\chi_{\nu+1}^2/2}}{2^{(\nu+1)/2}\Gamma\left(\nu/2\right)\Gamma\left(1/2\right)}\text{d}\chi_{\nu+1}^2\int_0^1 z^{\nu/2-1}\left(1-z\right)^{1/2-1}\text{d}z.\end{aligned} \tag{9}$$

However, the integral on the right side of Equation (9) is the beta function, $B\left(\nu/2, 1/2\right)$, which is related to the gamma functions by [22],

$$B\left(\nu/2, 1/2\right) = \frac{\Gamma\left(\nu/2\right)\Gamma\left(1/2\right)}{\Gamma\left(\left(\nu+1\right)/2\right)}. \tag{10}$$

By inserting Equation (10) into Equation (9), simplifying, and comparing with the left side of Equation (7), one obtains

$$\text{Pdf}\left(\chi_{\nu+1}^2 \mid \nu+1\right) = \frac{\left(\chi_{\nu+1}^2\right)^{(\nu+1)/2-1}e^{-\chi_{\nu+1}^2/2}}{2^{(\nu+1)/2}\Gamma\left(\left(\nu+1\right)/2\right)}, \tag{11}$$

which is the Pdf given by Equation (3) for $\nu+1$ degrees of freedom and it proves Equation (3) by induction.

By substituting $\varphi_i^2 = \left((x_i - \mu_i)/\sigma_i\right)^2$, Equation (2) becomes

$$\chi_\nu^2 = \sum_{i=1}^n \varphi_i^2 \tag{12}$$

The sum of independent random variables φ_i^2 is called a convolution and the joint distribution function for χ_ν^2 can be obtained by calculating an n-dimensional convolution integral. Exploring the properties of this convolution leads to simplifications, which have been used in the literature. By convoluting two gamma distributions $\chi_1^2 \equiv \varphi_i^2$ from Equation (5) and using the theorem that the convolution of two gammas is also a gamma, one obtains $\text{gamma}\left(\chi_2^2 \mid 2/2, 2\right)$ [9]. By continuing this process of convoluting with χ_1^2, it is easy to infer that the full convolution is equal to $\text{gamma}\left(\chi_\nu^2 \mid \nu/2, 2\right)$, where $\nu = n$, which the CSD given by Equation (3). This provides a simplified derivation of CSD using convolution.

Another simplified derivation of CSD uses the theorem that the Mgf of convolution is a product of individual Mgfs [10]. Thus, by calculating Mfg of χ_1^2 from Equation (5) and taking it to the nth power, one obtains the Mgf for χ_ν^2, where $\nu = n$. One can also calculate the Mgf of the gamma distribution and infer from a comparison that the CSD in Equation (3) is a special case of the gamma distribution [10].

In this work we provide yet another simplified derivation of the CSD using Laplace transform [23]. The Laplace transform of Equation (5) is equal to

$$\int_0^\infty \frac{\left(\chi_1^2\right)^{1/2-1} \mathrm{e}^{-\chi_1^2/2}}{2^{1/2}\Gamma\left(1/2\right)} \mathrm{e}^{-s\chi_1^2} \mathrm{d}\chi_1^2 = \left(\frac{1/2}{s+1/2}\right)^{1/2}. \tag{13}$$

Subsequently, we use a theorem that the Laplace transform of a nth convolution is a product of the individual transforms, *i.e.* $\left(\frac{1/2}{s+1/2}\right)^{n/2}$. By abbreviating $u = \chi_n^2$, the inverse Laplace transform results in the Pdf of u,

$$\text{Pdf}\left(u \mid n\right) = \frac{1}{2\pi i} \oint \left(\frac{1/2}{s+1/2}\right)^{n/2} \mathrm{e}^{su} \mathrm{d}s = \frac{1}{2^{n/2}} \frac{1}{2\pi i} \oint \frac{\mathrm{e}^{su}}{\left(s+1/2\right)^{n/2}} \mathrm{d}s. \tag{14}$$

To calculate the contour integral in Equation (14), we start with the Cauchy integration formula for an analytic function $f\left(s\right)$ of a complex variable s having a simple pole at s_0 [24]:

$$f\left(s_0\right) = \frac{1}{2\pi i} \oint \frac{f\left(s\right)}{s - s_0} \mathrm{d}s. \tag{15}$$

The $k-1$ times differentiation of Equation (15), where the differentiation can be of an integer or a fractional order [25], results in:

$$f^{(k-1)}\left(s_0\right) = \frac{\Gamma\left(k\right)}{2\pi i} \oint \frac{f\left(s\right)}{\left(s - s_0\right)^k} \mathrm{d}s. \tag{16}$$

By comparing Equation (14) to Equation (16), we infer that $f\left(s\right) = \mathrm{e}^{su}$, $s_0 = -1/2$, and $k = n/2$. By inserting these variables to Equation (16) and plugging it into Equation (14), we obtain:

$$\text{Pdf}\left(u \mid n\right) = \frac{1}{2^{n/2}\Gamma\left(n/2\right)} \left(\frac{\mathrm{d}^{n/2-1}}{\mathrm{d}s^{n/2-1}} \mathrm{e}^{su}\right)_{s=-1/2} = \frac{u^{n/2-1}\mathrm{e}^{-u/2}}{2^{n/2}\Gamma\left(n/2\right)}, \tag{17}$$

which is the CSD given by Equation (3) for $\nu = n$ and $\chi_n^2 = u$.

Another simplified derivation of the CSD uses the Bayesian inference and it is not related to the convolutions described above [11]. It uses a normal likelihood function for multiple samples. It also uses the transformational prior distributions: $\propto 1/\sigma$ for scale parameter σ and a constant for translation parameter μ [26]. Marginalizing the joint distribution (μ, σ) over μ results in the CSD, whereas marginalizing over σ results in the *t*-distribution [27].

In Section 5, we summarize the advantages and disadvantages of the simplified derivation methods of CSD described in this section.

3. Test Statistics

Several models for the CST statistics can be derived from the general Equation (2). For the expected value, we can use either the sample mean $\bar{x}$ or the population mean μ, whereas for the standard deviation we can use either individual standard deviations σ_i or the sample standard deviation σ_x. We do not know the population standard deviation for the data described in Section 4. Model test statistics $\sum\left((x_i - \bar{x})/\sigma_x\right)^2$ is always equal to $n-1$ and thus not useful. However, the model test statistics $\sum\left((x_i - \bar{x})/\sigma_i\right)^2$ can be used to test the variance. Other possibilities are to test for both the variance and location by employing model test statistics $\sum\left((x_i - \mu)/\sigma_i\right)^2$ or $\sum\left((x_i - \mu)/\sigma_x\right)^2$, if the population mean is known which is the case for the data in Section 4.

For the *t*-test we perform a standard one-sample test, where we calculate *t* variable as $(\bar{x} - \mu)/\left(\sigma_x/\sqrt{n}\right)$. The *t*-test is the location test. The results of all these test models using radioactivity data are presented in Section 4.

4. Chi-Square- and *t*-Test for Radioactivity Detection in Drinking Water

The most convenient method of measuring GA and GB radioactivity in drinking water is by gas proportional counting [28]. In this method, a given quantity of water is evaporated with nitric acid onto a stainless-steel planchet and dried, leaving a residue containing any radioactivity. The planchet is then counted on a gas proportional detector. Alpha and beta particles are counted simultaneously, and they are differentiated by much larger ionization caused by the former.

As stated in Section 1, this method must be able to determine GA and GB at the DL, to be verified by the CST [20] using a minimum of seven samples. EPA recommends a right-tail (RT) CST at 99% Confidence Level (CL), or 0.01 significance. To accomplish this, $n = 9$ samples of community drinking water were spiked with ^{230}Th and ^{90}Sr/^{90}Y radionuclides providing alpha and beta radioactivity, respectively. The spiking activities (*i.e.* the expected μ) were: 2.9888 ± 0.0402 pCi/L for alpha and 4.1860 ± 0.0549 pCi/L for beta, close to the required DL values. The values of spiking activities and their uncertainties were obtained from the standards traceable to the National Institute of Standards and Technology (NIST). Then the experimental procedure was followed, and the meas-

ured GA and GB activities x_i are depicted as points in Figure 1 and Figure 2, respectively.

Also shown in Figure 1 and Figure 2 are the individual standard deviations σ_i, depicted as vertical lines. These standard uncertainties are propagated, including the Poisson statistics of radioactivity counting and background subtraction, uncertainties of the detector efficiency, cross-talk between alpha and beta particles, as well as solution-pipetting uncertainties. Therefore, they are slightly different for different samples.

The GA results are described first. The sample average for GA is given by $\bar{x} = 3.0951$ pCi/L (red horizontal thick line) which is close to the expected μ (green horizontal thick line) as seen in Figure 1. The sample standard deviation is given by $\sigma_x = 0.7000$ pCi/L. The results of the variance test, as defined in Section 3, are given in column 3 of Table 1. The number of the degrees of freedom is $\nu = 8$ because one constraint is from calculating the mean. The observed χ^2 statistics is equal to 14.0 for gross alpha. The right-tail (RT) and left-tail (LT) χ^2 are calculated from the CSD at 0.01 significance each. Since $1.6 < 14.0 < 20.1$, each tail test passes at 0.01 significance and two-tail (2T) test passes at 0.02 significance. Then, the two combined variance/location tests, as defined in Section 3 are given in columns 4 and 5 using σ_i and σ_x, respectively. $\nu = n = 9$ in these cases, because there are no constraints. They both pass for GA.

The *t*-test statistics is calculated as described in Section 3 resulting in 0.45 for GA, as given in column 6 in Table 1. The RT probability of 0.33 and 2T probability of 0.66 are larger than 0.01 and 0.02, respectively, ensuring the passage of the location *t*-test.

The gross beta activities plotted in Figure 2, with the mean $\bar{x} = 5.1274$ pCi/L (red horizontal thick line) and $\sigma_x = 0.3050$ pCi/L differ significantly from the expected μ (green horizontal thick line) beyond the observed uncertainties. That fact did not affect the variance test which passed for GB (column 3 in Table 1). However, the observed χ^2 of 43.1 and 93.7 exceed the calculated RT χ^2 of 21.7 (columns 4 and 5 in Table 1), therefore the combined variance/location tests failed. This failure is supported by the *t*-test, where the high $t = 9.26$ (column 6) resulted in very low values of the RT and 2T probabilities (columns 7 and 8) and failures of the test for GB.

To elucidate the reasons for failure of the GB CST and *t*-test, fifteen non-spiked Method Blank (MB) community water samples were prepared and measured. The average GA activity was below detection; however, the average GB was 0.8121 ± 0.2801 pCi/L. This MB was then subtracted from the spiked GB results and the corrected GB activities are plotted in Figure 3. The mean of the corrected GB is $\bar{x} = 4.3153$ pCi/L ($\sigma_x = 0.3050$ pCi/L), very close to the value for spiked radioactivity. The corrected observed χ^2 are now 2.7, 3.2 and 9.6 (columns 3, 4, and 5 in Table 1) ensuring the passage of the three CSTs. This is supported by the passage of the *t*-test also (columns 6, 7, and 8).

Table 1. The results of χ^2- and *t*-tests. Abbreviations: RT right-tail, LT left-tail, 2T two-tail. Significance is 0.01 for each tail.

1	2	3	4	5	6	7	8
		χ^2-test				*t*-test	
Experiment, reference	Parameter	Variance, σ_i	Variance and location, σ_i	Variance and location, σ_x		Location	
	Deg free	8	9	9	8		
	Calc RT	20.1	21.7	21.7			
	Calc LT	1.6	2.1	2.1	*t*	RT prob	2T prob
Gross Alpha, **Figure 1**	Observed	14.0	13.4	8.2	0.45	0.33	0.66
	Test result	Passed	Passed	Passed		Passed	Passed
Gross Beta, **Figure 2**	Observed	3.8	43.1	93.7	9.26	7.5E−06	1.5E−05
	Test result	Passed	Failed	Failed		Failed	Failed
Gross Beta-MB subtracted, **Figure 3**	Observed	2.7	3.2	9.6	1.27	0.12	0.24
	Test result	Passed	Passed	Passed		Passed	Passed

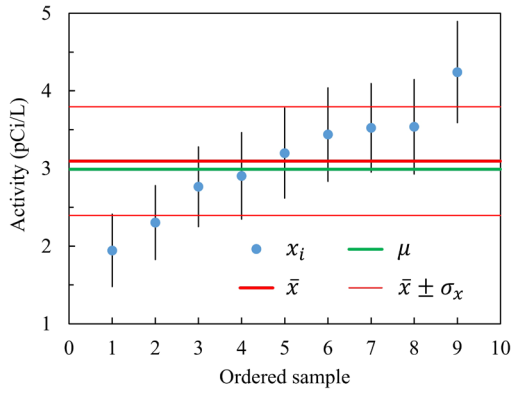


Figure 1. Gross alpha (points) ordered according to the increased activity.

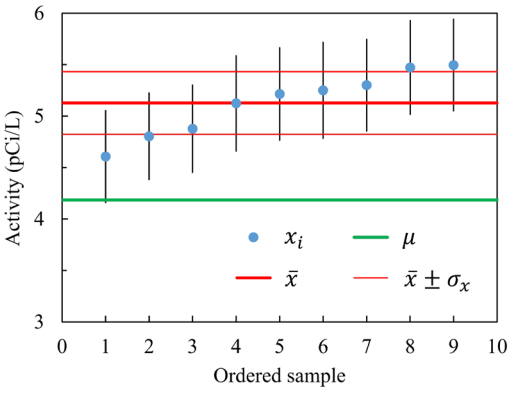


Figure 2. Gross beta (points) ordered according to the increased activity.

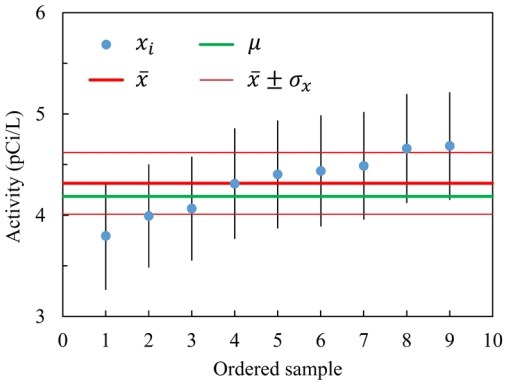


Figure 3. Gross beta (points) corrected for method blank and ordered according to the increased activity.

The reasons for the elevated GB in MB of community drinking water were investigated. Ten L of water were evaporated to 50 mL and measured using precise gamma-ray spectrometry [29]. It was determined that the concentration of the beta/gamma emitter, ^{40}K was 0.6926 ± 0.0790 pCi/L. It was also possible to identify several beta/gamma progenies of the ^{238}U series: ^{234}Th, ^{214}Pb, ^{214}Bi, and ^{210}Pb, as well as those from the ^{232}Th series: ^{228}Ac, ^{212}Pb, and ^{208}Tl. The combined activity of the beta/gamma progeny was 0.1513 ± 0.0672 pCi/L. Therefore, the sum of ^{40}K and beta/gamma progeny was 0.8440 ± 0.1037 pCi/L. The latter is consistent with the GB activity of 0.8121 ± 0.2801 pCi/L from the MB measurement to within the measured uncertainties. Also associated with the decay of ^{238}U and ^{232}Th is their alpha activity plus alpha progeny of similar activity to that of the beta/gamma progeny. This alpha activity could not have been detected by gamma spectrometry and was below the detection by GA in the MB measurement. However, the fact that GA of 3.0951 pCi/L is slightly higher than the expected 2.9888 pCi/L is an indication of that. Unlike in the case of beta activity, the small alpha progeny activity did not affect the CST or *t*-test. It should be noted that this level of naturally present radioactivity in the community water is much below the MCL, and thus poses small risk to the population.

5. Summary and Conclusions

We have described five simplified methods of deriving the chi-square distribution. Three of them: by convolution, moment generating function, and Bayesian inference are described in the literature and have been outlined here for comparison. The simplest of them seems to be the convolution method. It only uses the substitution from the normal distribution to a chi-square variable and requires a calculation of a single convolution integral on the above. It infers the form of multiple convolution on gamma distribution leading to the chi-square distribution. The moment generating function method of derivation is more advanced as it requires the knowledge of the moment generating function and the

gamma distribution. The Bayesian inference method requires the knowledge about likelihood function and prior probabilities but does not require the knowledge about the gamma distribution.

In this work, we have proposed two new methods for derivation of the chi-square distribution: by induction and by Laplace transform. The method of induction uses operational calculus with only a single integral leading to beta function. The proposed derivation applies modern formalism and seems to be simpler than the original derivation by Helmert as early as in 1876. A disadvantage of the induction method is that it requires a prior knowledge of the chi-square distribution to perform induction on it. There is a significant advantage, however. All other methods require either no constraints in the data; *i.e.* the number of degrees of freedom must be equal to the number of observations, or one constraint in case of Bayesian inference. The induction method leaves any constraints intact by adding one induction step to the existing number of degrees of freedom. The proposed derivation method by Laplace transform is more advanced because it uses integration in the complex plane. The significant advantage of the Laplace transform, and the Bayes inference methods is that they do not require prior knowledge about the gamma distribution.

We have also described a unique application of the chi-square test to environmental science. In chi-square testing, it is important to delineate systematic effects from the random uncertainties. In this work, a systematic natural contamination of laboratory method blank caused the chi-square test for combined variance/location to fail; however, it did not affect the chi-square test for variance alone. After subtracting the systematic method blank, the chi-square variance/location test was shown to have passed. This was confirmed by the location *t*-test. It is also imperative to perform analysis of uncertainty. In this work, using either individual or sample standard deviations did not affect the variance/location chi-square test. While the chi-square test provides verification if a laboratory test method is adequate to monitor gross alpha and gross beta radioactivity in drinking water, the test statistics combining variance and location is more useful than the one based on the variance alone because it can identify systematic bias.

Acknowledgements

N. F. acknowledges partial support by the Questar III STEM Research Institute for Teachers of Science, Engineering, Mathematics, and Technology. K. N. acknowledges partial support by the US Food and Drug Administration under Grant 5U18FD005514-04. Thanks are due to J. Witmer for his valuable comments.

Conflicts of Interest

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

References

[1] Hill, T.L. (1986) An Introduction to Statistical Thermodynamics. Dover Publications, New York, 122.

[2] Satchler, G.R. (1990) Introduction to Nuclear Reactions. Oxford U. P., New York, 248. https://doi.org/10.1007/978-1-349-20531-8

[3] Lancaster, H.O. (1969) The Chi-Squared Distribution. J. Wiley & Sons, New York, Chap. 1.

[4] Gorroochurn, P. (2016) Classic Topics on the History of Modern Mathematical Statistics: From Laplace to More Recent Times. J. Wiley & Sons, Hoboken, Chap. 3. https://doi.org/10.1002/9781119127963

[5] Bienaymé, I.-J. (1852) Sur la Probabilité des Erreurs d'Aprés la Méthode des Moindres Carrés. *Liouville's Journal de Mathématiques Poures et Appliquées*, *Séries* 1, **17**, 33-78.

[6] Abbe, D.E. (1863) Ueber die Gesetzmässigkeit in der Vertheilung der Fehler bei Beobachtungsreihen. Dissertation, Jena.

[7] Helmert, F.R. (1876) Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers directer Beobachtungen gleicher Genauigkeit. *Astronomische Nachrichten*, **88**, 113-132. https://doi.org/10.1002/asna.18760880802

[8] Stuart, A. and Ord, K. (1994) Kendall's Advanced Theory of Statistics, Vol. 1, Distribution Theory. Arnold Hodder Headline Group, London, Chap. 11.

[9] Ross, S. (2006) A First Course in Probability. Pearson Prentice Hall, Upper Saddle River, Sec. 6.3.

[10] Berry, D.A. and Lindgren, B.W. (1996) Statistics: Theory and Methods. Wadsworth Publishing, Belmont, Sec. 5.12, 6.4.

[11] Gull, S.F. (1988) Bayesian Inductive Inference and Maximum Entropy. In: Erickson, G.J. and Smith, C.R., Eds., *Maximum-Entropy and Bayesian Methods in Science and Engineering*, Kluwer Academic, Dordrecht, 53-74. https://doi.org/10.1007/978-94-009-3049-0_4

[12] Helmert, F.R. (1876) Ueber die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit im Zusammenhange stehende Fragen. *Zeitschrift für Mathematik und Physik*, **21**, 192-218.

[13] Pearson, K. (1900) On the Criterion that a Given System of Deviations from the Probable in the Case of Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling. *Philosophical Magazine Series* 5, **50**, 157-175. https://doi.org/10.1080/14786440009463897

[14] Greenwood, P.E. and Nikulin, M.S. (1996) A Guide to Chi-Squared Testing. J. Wiley & Sons, New York, Sec. 3.18.

[15] Fisher, R.A. (1922) On the Interpretation of χ^2 from Contingency Tables and the Calculation of P. *Journal of the Royal Statistical Society A*, **85**, 87-94. https://doi.org/10.2307/2340521

[16] Evans, R.D. (1985) The Atomic Nucleus. Krieger Publishing, Malabar, Chap. 27.

[17] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995) Continuous Univariate Distributions, Vol. 2. J. Wiley & Sons, New York, Chap. 28.

[18] Eisenbud, M. and Gesell, T. (1997) Environmental Radioactivity from Natural, Industrial, and Military Sources. Academic Press, San Diego, Chap. 6. https://doi.org/10.1016/B978-012235154-9/50010-4

[19] Environmental Protection Agency (2000) 40 CFR Parts 9, 141, and 142 National Primary Drinking Water Regulations; Radionuclides; Final Rule. *Federal Register*, **65**, 76708-76752.

[20] EPA (2017) Procedure for Safe Drinking Water Act Program Detection Limits for Radionuclides. Report EPA 815-B-17-003, Cincinnati.

[21] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994) Continuous Univariate Distributions, Vol. 1. J. Wiley & Sons, New York, Chap. 12, 17, 18.

[22] Johnson, N.L., Kemp, A.W. and Kotz, S. (2005) Univariate Discrete Distributions. J. Wiley & Sons, Hoboken, Chap. 1. https://doi.org/10.1002/0471715816

[23] Margenau, H. and Murphy, G.M. (1976) The Mathematics of Physics and Chemistry. Krieger Publishing, Huntington, Sec. 8.5.

[24] Dettman, J.W. (1965) Applied Complex Variables. Dover Publications, New York, Sec. 3.6.

[25] Oldham, K.B. and Spanier, J. (1974) The Fractional Calculus: Theory and Applications of Differentiation and Integration to Arbitrary Order. Dover Publications, Mineola, Sec. 3.4.

[26] Jaynes, E.T. (2004) Probability Theory: The Logic of Science. Cambridge U. P., Cambridge, Chap. 12. https://doi.org/10.1017/CBO9780511790423

[27] Student (1908) The Probable Error of a Mean. *Biometrika*, **6**, 1-25. https://doi.org/10.1093/biomet/6.1.1

[28] Semkow, T.M. and Parekh, P.P. (2001) Principles of Gross Alpha and Beta Radioactivity Detection in Water. *Health Physics*, **81**, 567-574. https://doi.org/10.1097/00004032-200111000-00011

[29] Khan, A.J., Semkow, T.M., Beach, S.E., Haines, D.K., Bradt, C.J., Bari, A., Syed, U.-F., Torres, M., Marrantino, J., Kitto, M.E., Menia, T. and Fielman, E. (2014) Application of Low-Background Gamma-Ray Spectrometry to Monitor Radioactivity in the Environment and Food. *Applied Radiation and Isotopes*, **90**, 251-257. https://doi.org/10.1016/j.apradiso.2014.04.011

Appendix

A.1. Glossary

CL: Confidence Level
CSD: Chi-Square Distribution
CST: Chi-Square Test
DL: Detection Limit for radionuclides
EPA: U.S. Environmental Protection Agency
GA: Gross Alpha Radioactivity
GB: Gross Beta Radioactivity
L: Liter
LT: Left Tail
MB: Method Blank
mBq: milli-Becquerel
MCL: Maximum Contaminant Level
MCLG: Maximum Contaminant Level Goal
Mgf: Moment generating function
mL: milli-Liter
mrem: milli-rem
NIST: National Institute of Standards and Technology
pCi: pico-Curie
Pdf: Probability density function
RT: Right Tail
SDWA: Safe Drinking Water Act
STEM: Science, Technology, Engineering and Mathematics
y: year
μSv: micro-Sievert
2T: Two Tail

A.2. Variables

a, b: parameters of the gamma distribution
B: beta function
E: expectation value
E_i: expected frequency
$f(s)$: analytic function
gamma: gamma distribution
i, k: indices
m: number of categories
n: number of observations
O_i: observed frequency
p: number of parameters for model distribution
s: complex variable
s_0: pole
t: t-test variable

Var: variance

x_i : normal random variable

$\overline{x}$: sample mean

u, z: substituted variables

Γ : gamma function

μ, μ_i : expected variable: population, individual

ν : number of degrees of freedom

$\sigma, \sigma_i, \sigma_x$: standard deviation, individual, sample

φ_i^2 : individual chi-square

$\chi^2, \chi_i^2, \chi_n^2, \chi_\nu^2$: chi-square, for i, n observations, ν degrees of freedom

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1800-1813
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Other Formulas for the Ree-Hoover and Mayer Weights of Families of 2-Connected Graphs

Amel Kaouche

Université de Moncton, Campus d'Edmundston, Edmundston, Canada
Email: amel.kaouche@umoncton.ca

How to cite this paper: Kaouche, A. (2019) Other Formulas for the Ree-Hoover and Mayer Weights of Families of 2-Connected Graphs. *Journal of Applied Mathematics and Physics*, **7**, 1800-1813.
https://doi.org/10.4236/jamp.2019.78123

Received: June 28, 2019
Accepted: August 16, 2019
Published: August 19, 2019

Abstract

We study graph weights which naturally occur in Mayer's theory and Ree-Hoover's theory for the virial expansion in the context of an imperfect gas. We pay particular attention to the Mayer weight and Ree-Hoover weight of a 2-connected graph in the case of the hard-core continuum gas in one dimension. These weights are calculated from signed volumes of convex polytopes associated with the graph. In the present paper, we use the method of graph homomorphisms, to develop other explicit formulas of Mayer weights and Ree-Hoover weights for infinite families of 2-connected graphs.

Keywords

Combinatorial, Mayer Weight, Statistical Mechanics, Ree-Hoover Weight, Graph Invariants, Virial Expansion

1. Introduction

Before discussing our subject, we first present some preliminary notions on the theory of graphs drawn from among others [1] [2] [3].

Preliminary Notions on the Theory of Graphs

Definition 1. *A* simple graph *g is formed of two sets: a non-empty finite set V, called the set of vertices of g, and a set E of pairs of vertices, called the set of edges of g. So we have* $E \subseteq \mathcal{P}_2(V)$ *with* $\mathcal{P}_2(V)$ *denotes all the parts of V with two elements. We often write* $g = (V, E)$.

Definition 2. *A* subgraph *h of a graph* $g = (V, E)$ *is a graph of the form* $h = (V_0, E_0)$, *such that* $V_0 \subseteq V$ *and* $E_0 = \mathcal{P}_2(V_0) \cap E$.

Definition 3. *An* over graph *g of a graph* $h = (V, E)$ *is a graph of the form* $g = (V_1, E_1)$, *such that* $V \subseteq V_1$ *and* $E = \mathcal{P}_2(V) \cap E_1$.

In the present work it will be useful to identify a graph with all of its edges, that is to say $g \subseteq \mathcal{P}_2(V)$.

Definition 4. *In a simple graph* $g=(V,E)$, *a* chain *c is a finite sequence of vertices,* $v_0, v_1, \cdots, v_m$, *such that for all* $0 \le i < m$, $\{v_i, v_{i+1}\} \in E$. *We write* $c = [v_0, v_1, \cdots, v_m]$.

Definition 5. *A graph* $g=(V,E)$ *is* connected *if* $\forall v, w \in V$, *there is a chain from v to w.*

Any graph breaks down uniquely as a disjoint union of connected graphs.

Definition 6. *On the set V of the vertices of the simple graph* $g=(V,E)$, *we define the relation of equivalence:* $v \sim w \Leftrightarrow$ *there is a chain v to w in g. Let* $V_1, V_2, \cdots, V_k$ *the equivalence classes of* $\sim$ *and let's say, for* $1 \le i \le k$, $g_i = g_{V_i}$, *the subgraph of g generated by* V_i. *These simple graphs* g_i, *that we call the* connected components *of g, are related* (see **Figure 1** *with connected components are circled*).

Definition 7. *A* cutpoint (or articulation point) *of a connected graph c is a vertex of c whose removal yields a disconnected graph.*

Definition 8. *A connected graph is called* 2*-connected if it has no cutpoint* (see **Figure 2**).

In the present paper, we study *Graph weights* in the context of a non-ideal gas in a vessel $V \subseteq \mathbb{R}^d$. In this case, the *Second Mayer weight* $w_M(c)$ of a connected graph c, over the set $[n] = \{1, 2, \cdots, n\}$ of vertices, is defined by (see [1] [4] [5] [6])

$$w_M(c) = \int_{(\mathbb{R}^d)^{n-1}} \prod_{\{i,j\} \in c} f\left(\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|\right) \mathrm{d}\boldsymbol{x}_1 \cdots \mathrm{d}\boldsymbol{x}_{n-1}, \quad \boldsymbol{x}_n = 0, \tag{1}$$

where $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ are variables in $\mathbb{R}^d$ representing the positions of n particles in V ($V \to \infty$), the value $\boldsymbol{x}_n = 0$ being arbitrarily fixed, and where $f = f(r)$ is a real-valued function associated with the pairwise interaction potential of the particles, see [6] [7].

Let $\mathcal{C}[n]$ be the set of connected graphs over $[n]$. The total sum of weights of connected graphs over $[n]$ is denoted by

$$\left|\mathcal{C}[n]\right|_{w_M} = \sum_{c \in \mathcal{C}[n]} w_M(c). \tag{2}$$

The interest of this sequence in statistical mechanics comes from the fact that the pressure P of the system is given by its exponential generating function as follows (see [6]):

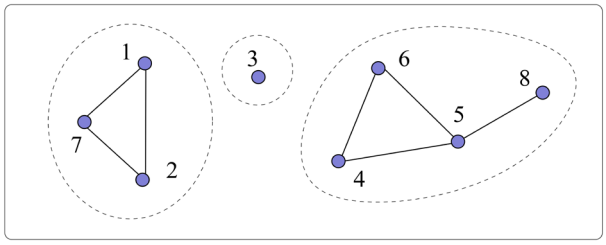


Figure 1. A simple graph and its connected components.

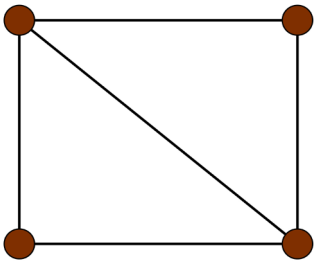


Figure 2. A 2-connected graph.

$$\frac{P}{kT} = \mathcal{C}_{w_M}(z) = \sum_{n\geq 1} \left|\mathcal{C}[n]\right|_{w_M} \frac{z^n}{n!}, \tag{3}$$

where k is a constant, T is the temperature, and z is a variable called the *fugacity* or the *activity* of the system.

It is known that the weight w_M is multiplicative over 2-connected components so that in order to compute the weights $w_M(c)$ of the connected graphs $c \in \mathcal{C}[n]$, it is sufficient to compute the weights $w_M(b)$ for 2-connected graphs $b \in \mathcal{B}[n]$ ($\mathcal{B}$ for *blocks*). The Mayer weight appear in the so-called *virial expansion* proposed by Kamerlingh Onnes in 1901

$$\frac{P}{kT} = \rho + \beta_2 \rho^2 + \beta_3 \rho^3 + \cdots, \tag{4}$$

where ρ is the density. Indeed, it can be shown that

$$\beta_n = \frac{1-n}{n!} \left|\mathcal{B}[n]\right|_{w_M}, \tag{5}$$

where $\mathcal{B}[n]$ denote the set of 2-connected graphs over $[n]$ and $\left|\mathcal{B}[n]\right|_{w_M}$ is the total sum of weights of 2-connected graphs over $[n]$. In order to compute this expansion numerically, Ree and Hoover [8] introduced a modified weight denoted by $w_{RH}(b)$, for 2-connected graphs b, which greatly simplifies the computations. It is defined by

$$w_{RH}(b) = \int_{\left(\mathbb{R}^d\right)^{n-1}} \prod_{\{i,j\}\in b} f\left(\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|\right) \prod_{\{i,j\}\notin b} \overline{f}\left(\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|\right) \mathrm{d}\boldsymbol{x}_1 \cdots \mathrm{d}\boldsymbol{x}_{n-1}, \quad \boldsymbol{x}_n = 0, \tag{6}$$

where $\overline{f}(r) = 1 + f(r)$. Using this new weight, Ree and Hoover [8] [9] [10] and later Clisby and McCoy [11] [12] [13] have computed the virial coefficients β_n, for n up to 10, in dimensions $d \leq 8$, in the case of the hard-core continuum gas, that is when the interaction is given by

$$f(r) = -\chi(r<1), \quad \overline{f}(r) = \chi(r \geq 1), \tag{7}$$

where χ denote the characteristic function ($\chi(P) = 1$, if P is true and 0, otherwise).

The main goal of the present paper is to give new explicit formulas for the Mayer and Ree-Hoover weights of certain infinite families of graphs in the context of the hard core continuum gas, defined by (7), in dimension $d = 1$. The values $w_M(c)$ and $w_{RH}(c)$ for all 2-connected graphs c of size at most 8 are given in [1] [14].

In Section 2, we look at the case of the hard-core continuum gas in one dimension in which the Mayer weight turns out to be a signed volume of a convex polytope $\mathcal{P}(c)$ naturally associated with the graph c. A decomposition of the polytope $\mathcal{P}(c)$ into a certain number of simplices is utilised. This method was introduced in [6] and was adapted in [1] [5] to the context of Ree-Hoover weights and is called the method of graph homomorphisms. The explicit computation of Mayer or Ree-Hoover weights of particular graphs is very challenging in general and have been made for only certain specific families of graphs (see [4] [5] [6] [15] [16] [17] [18]). In the present paper we extend this list to include other graphs. We give new explicit formulas of the Ree-Hoover weight of these graphs in Section 3. Section 4 is devoted to the explicit computation of their Mayer weight. The following conventions are used in the present paper: Each graph g is identified with its set of edges. So that, $\{i,j\}\in g$ means that $\{i,j\}$ is an edge in g between vertex i and vertex j. The number of edges in g is denoted $e(g)$. If e is an edge of g (*i.e.* $e\in g$), $g\setminus e$ denotes the graph obtained from g by removing the edge e. If b and d are graphs, $b\subseteq d$ means that b is a subgraph of d. The complete graph on the vertex set $[n]=\{1,2,\cdots,n\}$ is denoted by K_n. The complementary graph of a subgraph $g\subseteq K_n$ is the graph $\overline{g}=K_n\setminus g$.

An important rewriting of the virial coefficients was performed by Ree and Hoover [8] [9] by introducing the function

$$\overline{f}(r)=1+f(r) \tag{8}$$

and defining a new weight (denoted here by $w_{RH}(b)$) for 2-connected graphs b, by (9)

$$w_{RH}(b)=\int_{\left(\mathbb{R}^d\right)^{n-1}}\prod_{\{i,j\}\in b}f\left(\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|\right)\prod_{\{i,j\}\notin b}\overline{f}\left(\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|\right)\mathrm{d}\boldsymbol{x}_1\cdots\mathrm{d}\boldsymbol{x}_{n-1},\quad \boldsymbol{x}_n=0, \tag{9}$$

and then expanding each weight $w_M(b)$ by substituting $1=\overline{f}-f$ for pairs of vertices not connected by edges.

In [1], we gived explicit linear relations expressing the Ree-Hoover weights in terms of the Mayer weights and vice versa: For a 2-connected graph b, we have

$$w_{RH}(b)=\sum_{b\subseteq d\subseteq K_n}w_M(d), \tag{10}$$

$$w_M(b)=\sum_{b\subseteq d\subseteq K_n}(-1)^{e(d)-e(b)}w_{RH}(d). \tag{11}$$

So that the virial coefficient can be rewritten in the form

$$\beta_n=\frac{1-n}{n!}\sum_{b\in\mathcal{B}[n]}a_n(b)w_{RH}(b), \tag{12}$$

for appropriate coefficients $a_n(b)$ called the *star content* of the graph b. The importance of (1.12) is due to the fact that $a_n(b)=0$ or $w_{RH}(b)=0$ for many graphs b. This greatly simplifies the computation of β_n.

Using the definition of the Ree-Hoover weight, we have

$$w_{RH}(K_n)=w_M(K_n),\quad n\geq 2. \tag{13}$$

2. Hard-Core Continuum Gas in One Dimension

Consider n hard particles of diameter 1 on a line segment. The *hard-core* constraint translates into the interaction potential φ, with $\varphi(r)=\infty$, if $r<1$, and $\varphi(r)=0$, if $r\geq 1$, and the Mayer function f and the Ree-Hoover function $\overline{f}$ are given by (7). Hence, we can write the Mayer weight function $w_M(c)$ of a connected graph c as

$$w_M(c)=(-1)^{e(c)}\int_{\mathbb{R}^{n-1}}\prod_{\{i,j\}\in c}\chi\left(\left|x_i-x_j\right|<1\right)\mathrm{d}x_1\cdots\mathrm{d}x_{n-1},\ x_n=0, \tag{14}$$

and the Ree-Hoover's weight function $w_{RH}(c)$ of a 2-connected graph c as

$$w_{RH}(c)=(-1)^{e(c)}\int_{\mathbb{R}^{n-1}}\prod_{\{i,j\}\in c}\chi\left(\left|x_i-x_j\right|<1\right)\prod_{\{i,j\}\notin c}\chi\left(\left|x_i-x_j\right|>1\right)\mathrm{d}x_1\cdots\mathrm{d}x_{n-1}, \tag{15}$$

with $x_n=0$ and where $e(c)$ is the number of edges of c. Note that $w_M(c)=(-1)^{e(c)}\operatorname{Vol}(\mathcal{P}(c))$, where $\mathcal{P}(c)$ is the polytope defined by

$$\mathcal{P}(c)=\left\{X\in\mathbb{R}^n\middle|x_n=0,\left|x_i-x_j\right|<1\ \forall\{i,j\}\in c\right\}\subseteq\mathbb{R}^{n-1}\times\{0\}\subseteq\mathbb{R}^n,$$

where $X=(x_1,\cdots,x_n)$. Similarly, $w_{RH}(c)=(-1)^{e(c)}\operatorname{Vol}(\mathcal{P}_{RH}(c))$, where $\mathcal{P}_{RH}(c)$ is the union of polytopes defined by

$$\mathcal{P}_{RH}(c)=\left\{X\in\mathbb{R}^n\middle|x_n=0,\left|x_i-x_j\right|<1\ \forall\{i,j\}\in c,\left|x_i-x_j\right|>1\ \forall\{i,j\}\in\overline{c}\right\}.$$

Graph Homomorphisms

The method of graph homomorphisms was introduced in [6] for the calculation of the Mayer weight $w_M(b)$ of a 2-connected graph b in the context of hard-core continuum gases in one dimension and was fited in [5] to the context of Ree-Hoover weights. Since $w_M(b)=(-1)^{e(b)}\operatorname{Vol}(\mathcal{P}(b))$, the calculation of $w_M(b)$ is reduced to the calculation of the volume of the polytope $\mathcal{P}(b)$ associated to b. In order to compute this volume, the polytope $\mathcal{P}(b)$ is decomposed into $\nu(b)$ simplices which are all of volume $1/(n-1)!$ and we will have $\operatorname{Vol}(\mathcal{P}(b))=\nu(b)/(n-1)!$. Each simplice is represented by a diagram associated to the integral parts and the relative positions of the fractional parts of the coordinates $x_1,\cdots,x_n$ of points $X\in\mathcal{P}(b)$.

More specifically, to each real number x, they associate his fractional representation, which is a pair (ξ_x,h_x), where $h_x=\lfloor x\rfloor$ is the integral part of x and $\xi_x=x-h_x$ is the (positive) fractional part of x, so that $x=\xi_x+h_x$. Then, for $x\neq y$, the condition $|x-y|<1$ translates into "assuming $\xi_x<\xi_y$, then $h_x=h_y$ or $h_x=h_y+1$". It mean that the slope of the line segment between the points (ξ_x,h_x) and (ξ_y,h_y) in the plane should be either null or negative. Let b a 2-connected graph with vertex set $[n]$, and let $X=(x_1,\cdots,x_n)$ be a point in the polytope $\mathcal{P}(b)$. Let's write (ξ_i,h_i) for the fractional representation of the coordinate x_i of X. For $x_n=0$, it will be convenient to use the special representation $\xi_n=1.0$ and $h_n=-1$. Remarque that the volume of $\mathcal{P}(b)$ is unchanged by removing all hyperplanes $\{x_i-x_j=k\}$, for $k\in\mathbb{Z}$. in consequence, we can assume that all the fractional parts ξ_i are distinct. We get a subpolytope of $\mathcal{P}(b)$ by fixing the "heights" $h_1,h_2,\cdots,h_n$ as well as the relative positions

(total order) of the fractional parts $\xi_1, \xi_2, \cdots, \xi_n$. Let $h: V \to \mathbb{Z}$ denote the height function $i \mapsto h_i$ and $\beta: V \to [n]$ be the permutation of $[n]$ for which $\beta(i)$ gives the rank of ξ_i in this total order with $\beta(n) = n$. Explicitly, each simplex $\mathcal{P}(h, \beta)$ can be written as

$$\mathcal{P}(h,\beta) = \left\{ (h_1 + \xi_1, \cdots, h_{n-1} + \xi_{n-1}, 0) \middle| 0 < \xi_{\beta^{-1}(1)} < \cdots < \xi_{\beta^{-1}(n-1)} < 1 \right\} \quad (16)$$

and it is shown in [1] that each such simplex is affine-equivalent to the standard simplex

$$\mathcal{P}(0, \mathrm{id}) = \left\{ (\xi_1, \xi_2, \cdots, \xi_{n-1}, 0) \middle| 0 < \xi_1 < \xi_2 < \cdots < \xi_{n-1} < 1 \right\}$$

of volume $1/(n-1)!$.

Note that the simplices (16) are disjoint and each such simplex can be characterized by its centre of gravity

$$X_{h,\beta} = \left(h_1 + \frac{\beta(1)}{n}, h_2 + \frac{\beta(2)}{n}, \cdots, h_{n-1} + \frac{\beta(n-1)}{n}, 0 \right).$$

Note also that when there are no restrictions on h and β, the union of the closed simplices $\overline{\mathcal{P}(h,\beta)}$ coincides with the whole configurations space $\mathbb{R}^{n-1} \times \{0\}$.

Using the fractional coordinates to represent the center of gravity $X_{h,\beta}$ of the simplex $\mathcal{P}(h,\beta)$, and drawing a line segment form $x_i = (h_i, \xi_i)$ and $x_j = (h_j, \xi_j)$ for each edge $\{i, j\}$ of the graph b, we get a configuration in the plane which is an homomorphic image of b which represents the subpolytope $\mathcal{P}(h,\beta)$. The above content is summarized in to the form of a proposition:

Proposition 1. ([6]). *Let b be a 2-connected graph with vertex set $V = [n]$ and consider a function $h: V \to \mathbb{Z}$ and a bijection $\beta: V \to [n]$ satisfying $\beta(n) = n$. Then the simplex $\mathcal{P}(h,\beta)$ corresponding to the pair (h,β) is contained in the polytope $\mathcal{P}(\beta)$ if and only if the following condition is satisfied*:

for any edge $\{i, j\}$ of b, $\beta(i) < \beta(j)$ implies $h_i = h_j$ or $h_i = h_j + 1$. (17)

Corollary 1. ([6]). *Let b be a 2-connected graph and let $\nu(b)$ be the number of pairs (h,β) such that the condition* (17) *is satisfied. Then the volume of the polytope $\mathcal{P}(b)$ is given by*

$$\mathrm{Vol}(\mathcal{P}(b)) = \nu(b)/(n-1)!. \quad (18)$$

Proposition 1 can be used to compute the weight of some families of graphs, since $w_M(b) = (-1)^{e(b)} \mathrm{Vol}(\mathcal{P}(b))$.

In a similar way we can adapt the above configurations to the context of the Ree-Hoover weight.

Proposition 2. ([5]). *Let b be a 2-connected graph with vertex set $V = [n]$ and consider a function $h: V \to \mathbb{Z}$ and a bijection $\beta: V \to [n]$ satisfying $\beta(n) = n$. Then the simplex $\mathcal{P}(h,\beta)$ corresponding to the pair (h,β) is contained in the polytope $\mathcal{P}_{RH}(b)$ if and only if the following conditions are satisfied*:

for any edge $\{i,j\}$ of b, $\beta(i)<\beta(j)$ implies $h_i = h_j$ or $h_i = h_j + 1$. (19)

for any edge $\{i,j\}$ of $\overline{b}$, $\beta(i)<\beta(j)$ implies $h_i \le h_j - 1$ or $h_i \ge h_j + 2$. (20)

Proposition 3. ([5]). *Let b be a 2-connected graph and let $v_{RH}(b)$ be the number of pairs (h,β) such that conditions* (19) *and* (20) *are satisfied. Then the volume of $\mathcal{P}_{RH}(b)$ is given by*

$$\mathrm{Vol}\left(\mathcal{P}_{RH}(b)\right) = v_{RH}(b)/(n-1)!. \tag{21}$$

3. Ree-Hoover Weight of New Families of Graphs

In this section, we give other explicit formulas for the Ree-Hoover weight for infinite families of 2-connected graphs. First, we use Ehrhart polynomials to conjectured these formulas from numerical values. We use the techniques of *graph homorphisms* in order to prove these formulas. The weights of 2-connected graphs b are given in absolute value $|w(b)|$, the sign being always equal to $(-1)^{e(b)}$.

Lemma 1. ([5]). *Suppose that g is a graph over $[n]$ and $i,j \in [n-1]$ are such that g does not contain the edge $\{n,i\}$ but contains the edges $\{i,j\}$ and $\{n,j\}$. In this case, any RH-configuration (h,β) (with $h_n = -1$, $\beta(n) = n$) satisfies either one of the following conditions:*

1) $h_i = 1$, $h_j = 0$ and $\beta(i) < \beta(j)$,
2) $h_i = -2$, $h_j = -1$ and $\beta(i) > \beta(j)$.

3.1. The Ree-Hoover Weight of the Graph $K_n \setminus ((C_4 \cdot\cdot S_2)\cdot S_1)$

Let $(C_4 \cdot\cdot S_2)\cdot S_k$ denote the graph obtained by identifying one vertex, with degree three, of the graph $(C_4 \cdot\cdot S_2)$ with a center of a k-star. See **Figure 3** for an example.

Let us start with the simple case $(C_4 \cdot\cdot S_2)\cdot S_1$.

Proposition 4. *For $n \ge 7$, we have*

$$\left|w_{RH}\left(K_n \setminus ((C_4 \cdot\cdot S_2)\cdot S_1)\right)\right| = \frac{12}{(n-1)(n-2)(n-3)(n-4)(n-5)}. \tag{22}$$

Proof. We can assume that the missing edges are $\{1,n\}$, $\{2,n\}$, $\{4,n\}$, $\{n,5\}$, $\{1,3\}$, $\{3,4\}$ and $\{2,3\}$ (see **Figure 4**).

According to Lemma 1 there are two possibilities for h:

- $h_1 = h_2 = h_4 = h_5 = 1$ and $h_n = -1$ and all other $h_i = 0$, so that $\beta(5) = 1$ and $\beta(3) = 2$ and $(\beta(1),\beta(2),\beta(4))$ must be a permutation of $\{3,4,5\}$.
- $h_1 = h_2 = h_4 = h_5 = -2$ and all other $h_i = -1$, so that $\beta(5) = n-1$ and $\beta(3) = n-2$ and $(\beta(1),\beta(2),\beta(4))$ must be a permutation of $\{n-3,n-4,n-5\}$.

In each case β can be extended in $(n-6)!$ ways, giving the possible relative positions of the $(n-6)$ x_i (see **Figure 5**). So, there are $2\cdot 3!(n-6)!$ RH-configurations (h,β).

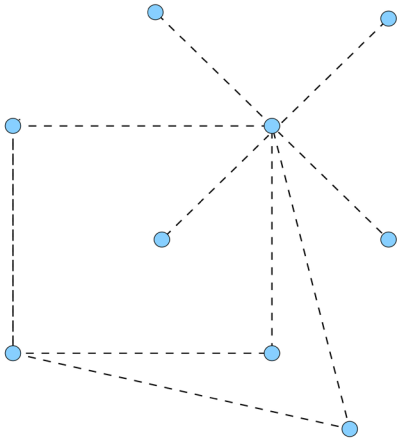


Figure 3. The graph $(C_4 \cdot\cdot S_2)\cdot S_4$.

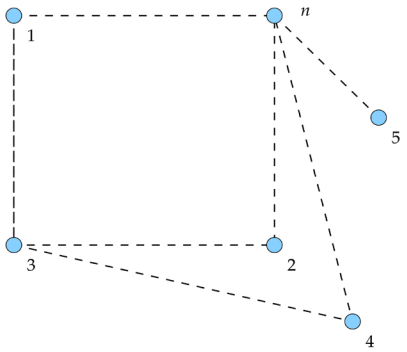


Figure 4. The graph $(C_4 \cdot\cdot S_2)\cdot S_1$.

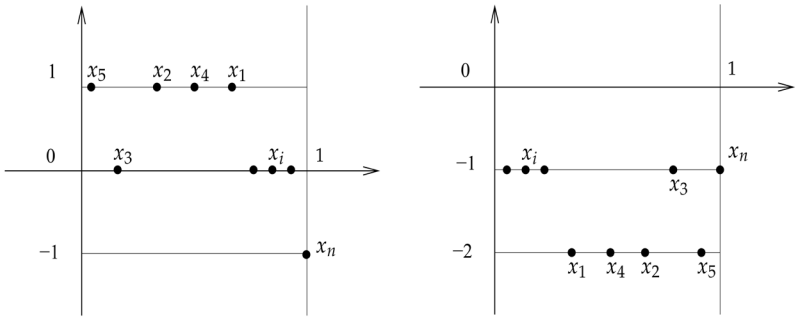


Figure 5. Fractional representation of a simplicial subpolytope of $\mathcal{P}_{RH}\left(K_n \setminus \left((C_4 \cdot\cdot S_2)\cdot S_1\right)\right)$.

3.2. The Ree-Hoover Weight of the Graph $K_n \setminus \left((C_4 \cdot\cdot S_2)\cdot S_k\right)$

In the general case we have:

Proposition 5. *For* $k \geq 1$, $n \geq k+6$, *we have*

$$\left|w_{RH}\left(K_n \setminus \left((C_4 \cdot\cdot S_2)\cdot S_k\right)\right)\right| = \frac{12k!}{(n-1)(n-2)\cdots(n-k-4)}. \tag{23}$$

Proof. We can assume that the missing edges are $\{1,n\}$, $\{2,n\}$, $\{4,n\}$,

$\{3,4\}$, $\{2,3\}$, $\{1,3\}$ and $\{5,n\}$, $\{6,n\}$, $\cdots$, $\{k+4,n\}$ (see **Figure 6**, for the case of $(C_4 \cdot\cdot S_2)\cdot S_2$).

According to Lemma 1 there are two possibilities for h:

- $h_1 = h_2 = h_4 = h_5 = \cdots = h_{k+4} = 1$ and $h_n = -1$ and all other $h_i = 0$, so that $(\beta(5), \beta(6), \beta(7), \cdots, \beta(k+4))$ must be a permutation of $\{1,2,3,\cdots,k\}$ and $\beta(3) = k+1$ and $(\beta(1), \beta(2), \beta(4))$ must be a permutation of $\{k+2, k+3, k+4\}$.
- $h_1 = h_2 = h_4 = h_5 = \cdots = h_{k+4} = -2$ and all other $h_i = -1$, so that $(\beta(5), \beta(6), \beta(7), \cdots, \beta(k+4))$ must be a permutation of $\{n-1, n-2, n-3, \cdots, n-k\}$ and $\beta(3) = n-k-1$ and $(\beta(1), \beta(2), \beta(4))$ must be a permutation of $\{n-k-2, n-k-3, n-k-4\}$.

In each case β can be extended in $(n-k-5)!$ ways, giving the possible relative positions of the $(n-k-5)$ x_i (see **Figure 7**, for the case of $S_2 \cdot C_4 \cdot\cdot S_2$). So, there are $2 \cdot 3! k! (n-k-5)!$ RH-configurations (h, β).

We need to use Propositions (6)-(10) to prove Mayer's weight formulas that will be presented in section 4.

Proposition 6. ([5]). For $n \geq 6$, we have

$$\left| w_{RH}\left(K_n \setminus C_4\right) \right| = \frac{8}{(n-1)(n-2)(n-3)}, \tag{24}$$

where C_4 is the unoriented cycle with 4 vertices.

Proposition 7. ([5]). For $k \geq 1$, $n \geq k+3$, we have

$$\left| w_{RH}\left(K_n \setminus S_k\right) \right| = \frac{2k!}{(n-1)(n-2)\cdots(n-k)}, \tag{25}$$

where S_k denote the k-star graph with vertex set $[k+1]$ and edge set $\{\{1,2\},\{1,3\},\cdots,\{1,k+1\}\}$, (see **Figure 8**, for the case of S_3).

Proposition 8. ([5]). For $j \geq k \geq 1$, $n \geq k+j+3$, we have

$$\left| w_{RH}\left(K_n \setminus \left(S_j - S_k\right)\right) \right| = \frac{2k!\, j!}{(n-1)(n-2)\cdots(n-(k+j+1))}, \tag{26}$$

where $S_j - S_k$ denote the graph obtained by joining with a new edge the centers of a j-star and of a k-star. See **Figure 9** for an example.

Proposition 9. ([5]). For $k \geq 1$, $n \geq k+5$, we have

$$\left| w_{RH}\left(K_n \setminus (C_4 \cdot S_k)\right) \right| = \frac{4k!}{(n-1)(n-2)\cdots(n-(k+3))}, \tag{27}$$

where $C_4 \cdot S_k$ denote the graph obtained by identifying one vertex of the graph C_4 with the center of a k-star. See **Figure 10** for an example.

Proposition 10. ([18]). For $n \geq 7$, we have

$$\left| w_{RH}\left(K_n \setminus (C_4 \cdot\cdot S_2)\right) \right| = \frac{24}{(n-1)(n-2)(n-3)(n-4)}, \tag{28}$$

where $(C_4 \cdot\cdot S_2)$ denote the graph obtained by identifying two non adjacent vertices of the graph C_4 (the unoriented cycle with 4 vertices) with the extremities of a 2-star (see **Figure 11**).

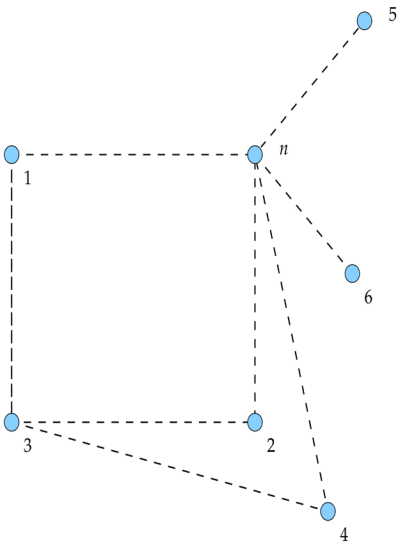


Figure 6. The graph $(C_4 \cdot\cdot S_2)\cdot S_2$.

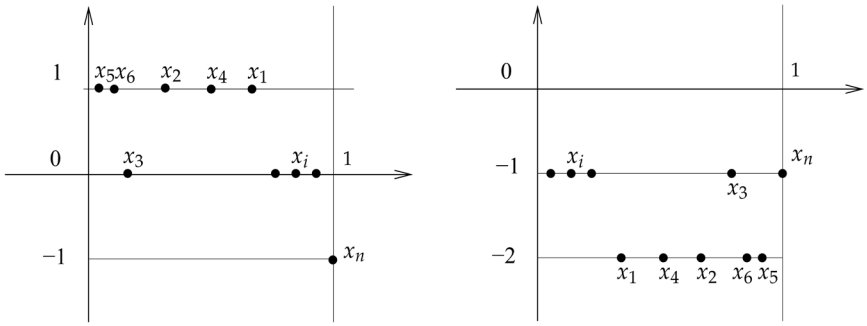


Figure 7. Fractional representation of a simplicial subpolytope of $\mathcal{P}_{RH}\left(K_n \setminus \left((C_4 \cdot\cdot S_2)\cdot S_2\right)\right)$.

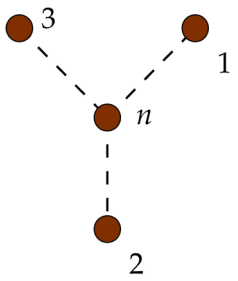


Figure 8. The graph S_3.

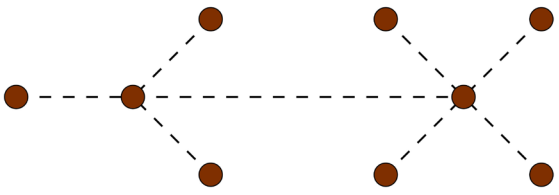


Figure 9. The graph $S_3 - S_4$.

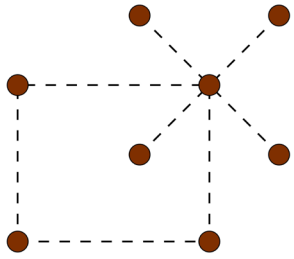


Figure 10. The graph $C_4 \cdot S_4$.

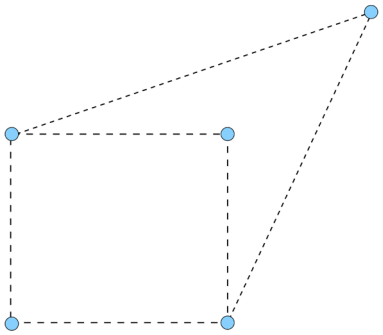


Figure 11. The graph $C_4 \cdot\cdot S_2$.

4. Mayer Weight of New Families of Graphs

Here are some of our results concerning new explicit formulas for the Mayer weight of the previous infinite families of graphs. In this case, the computation of the Mayer weight is more difficult. Instead of using the method of graph homomorphisms, we use the following formula

$$\left|w_M(b)\right| = \sum_{b \subseteq d \subseteq K_n} \left|w_{RH}(d)\right| \tag{29}$$

which is a consequence of (1.11) in the case of hard-core continuum gases in one dimension. Substituting $K_n \setminus g$ and $K_n \setminus k$ for b and d in (29), we have

$$\left|w_M(K_n \setminus g)\right| = \sum_{k \subseteq g} \left|w_{RH}(K_n \setminus k)\right| = \sum_{\tilde{h} \subseteq \tilde{g}} m\left(\tilde{h}, \tilde{g}\right) \left|w_{RH}(K_n \setminus h)\right|, \tag{30}$$

where $\tilde{g}$ denotes the unlabelled graph corresponding to g, $\tilde{h}$ runs through the unlabelled subgraphs of $\tilde{g}$ and $m\left(\tilde{h}, \tilde{g}\right)$ is the number of ways of obtaining $\tilde{h}$ by removing some edges in $\tilde{h}$. We obtain these multiplicities $m\left(\tilde{h}, \tilde{g}\right)$ by combinatorial arguments.

4.1. The Mayer Weight of the Graph $K_n \setminus \left(\left(C_4 \cdot\cdot S_2\right) \cdot S_1\right)$

Proposition 11. *For* $n \geq 7$, *we have*

$$\begin{aligned}\left|w_M\left(K_n \setminus \left(C_4 \cdot\cdot S_2\right) \cdot S_1\right)\right| = n + \frac{14}{n-1} + \frac{48}{(n-1)(n-2)} + \frac{114}{(n-1)\cdots(n-3)} \\ + \frac{156}{(n-1)\cdots(n-4)} + \frac{72}{(n-1)\cdots(n-5)}.\end{aligned} \tag{31}$$

Proof. The over graphs of $K_n \setminus (C_4 \cdot\cdot S_2)\cdot S_1$ whose Ree-Hoover weight is not zero are up to isomorphism of the form: $K_n \setminus C_4$, $K_n \setminus C_4 \cdot S_l$, $1 \le l \le 2$, $K_n \setminus C_4 \cdot\cdot S_2$, $K_n \setminus (C_4 \cdot\cdot S_2)\cdot S_1$, $K_n \setminus S_l$, $1 \le l \le 4$, $K_n \setminus (S_1 - S_l)$, $1 \le l \le 3$, and K_n. Their multiplicities are given by

$$\begin{aligned}
&\left|w_M\left(K_n \setminus (C_4 \cdot\cdot S_2)\cdot S_1\right)\right| \\
&= \left|w_{RH}\left(K_n\right)\right| + 7\left|w_{RH}\left(K_n \setminus S_1\right)\right| + 12\left|w_{RH}\left(K_n \setminus S_2\right)\right| + 5\left|w_{RH}\left(K_n \setminus S_3\right)\right| \\
&\quad + \left|w_{RH}\left(K_n \setminus S_4\right)\right| + 3\left|w_{RH}\left(K_n \setminus C_4\right)\right| + 15\left|w_{RH}\left(K_n \setminus (S_1 - S_1)\right)\right| \\
&\quad + 12\left|w_{RH}\left(K_n \setminus (S_1 - S_2)\right)\right| + 3\left|w_{RH}\left(K_n \setminus (S_1 - S_3)\right)\right| + \left|w_{RH}\left(K_n \setminus (C_4 \cdot\cdot S_2)\right)\right| \\
&\quad + 9\left|w_{RH}\left(K_n \setminus (C_4 \cdot S_1)\right)\right| + 3\left|w_{RH}\left(K_n \setminus (C_4 \cdot S_2)\right)\right| + \left|w_{RH}\left(K_n \setminus ((C_4 \cdot\cdot S_2)\cdot S_1)\right)\right|.
\end{aligned}$$

We conclude using Propositions (5) and (6)-(10).

4.2. The Mayer Weight of the Graph $K_n \setminus ((C_4 \cdot\cdot S_2)\cdot S_k)$

Proposition 12. *For* $k \ge 1$, $n \ge k+6$, $g_n = K_n \setminus ((C_4 \cdot\cdot S_2)\cdot S_k)$, *we have, with the usual convention* $\binom{k+1}{\ell} = 0$ *if* $\ell > k+1$,

$$\begin{aligned}
\left|w_M\left(g_n\right)\right| &= n + \frac{6}{(n-1)} + \frac{60}{(n-1)(n-2)(n-3)(n-4)} \\
&\quad + \frac{24}{(n-1)(n-2)} + \frac{60}{(n-1)(n-2)(n-3)} \\
&\quad + \sum_{l=1}^{k+3}\left[\binom{k+3}{l}\frac{2l!}{(n-1)\cdots(n-l)} + \binom{k+1}{l}\frac{12l!}{(n-1)\cdots(n-l-3)}\right] \\
&\quad + \sum_{l=1}^{k+2}\left[\binom{k}{l-2} + 2\binom{k}{l-1} + \binom{k}{l}\right]\left[\frac{6l!}{(n-1)\cdots(n-l-2)}\right] \\
&\quad + \sum_{l=1}^{k}\binom{k}{l}\left[\frac{12l!}{(n-1)\cdots(n-l-4)}\right].
\end{aligned}$$

Proof. The over graphs of $K_n \setminus ((C_4 \cdot\cdot S_2)\cdot S_k)$ whose Ree-Hoover weight is not zero are up to isomorphism of the form: $K_n \setminus S_l$, $1 \le l \le k+3$, $K_n \setminus (C_4 \cdot S_l)$, $1 \le l \le k+1$, $K_n \setminus (S_1 - S_l)$, $1 \le l \le k+2$, $K_n \setminus C_4$, $K_n \setminus (C_4 \cdot\cdot S_2)$, $K_n \setminus ((C_4 \cdot\cdot S_2)\cdot S_l)$, $1 \le l \le k$ and K_n. Their multiplicities are given by

$$\begin{aligned}
\left|w_M\left(g_n\right)\right| &= \left|w_{RH}\left(K_n\right)\right| + 3\left|w_{RH}\left(K_n \setminus S_1\right)\right| + 6\left|w_{RH}\left(K_n \setminus S_2\right)\right| + \left|w_{RH}\left(K_n \setminus S_3\right)\right| \\
&\quad + 3\left|w_{RH}\left(K_n \setminus C_4\right)\right| + 12\left|w_{RH}\left(K_n \setminus S_1 - S_1\right)\right| + 6\left|w_{RH}\left(K_n \setminus S_1 - S_2\right)\right| \\
&\quad + \left|w_{RH}\left(K_n \setminus (C_4 \cdot\cdot S_2)\right)\right| + 3\left|w_M\left(K_n \setminus (C_4 \cdot S_1)\right)\right| \\
&\quad + \sum_{l=1}^{k+1} 3\binom{k+1}{l}\left[\left|w_{RH}\left(K_n \setminus C_4 \cdot S_l\right)\right| + \binom{k}{l}\left|w_{RH}\left(K_n \setminus (C_4 \cdot\cdot S_2)\cdot S_l\right)\right|\right] \\
&\quad + \sum_{l=1}^{k+2} 3\left[\binom{k}{l} + 2\binom{k}{l-1} + \binom{k}{l-2}\right]\left|w_{RH}\left(K_n \setminus S_1 - S_l\right)\right|
\end{aligned}$$

$$+\sum_{l=1}^{k+3}\binom{k+3}{l}\left|w_{RH}\left(K_n \setminus S_l\right)\right|.$$

We conclude using Propositions (5) and (6)-(10).

5. Conclusion

The links between statistical mechanics and combinatorics are more and more numerous as we have seen in this work. In this paper, after recalling the Mayer and Ree-Hoover theory, we presented in Section 2 the method of graph homomorphisms and we have mainly placed ourselves in the context of hard-core continuum gas in one dimension. From various tables that we constructed giving numerical values of Mayer and Ree-Hoover weights of all 2-connected graphs up to size 8, we conjectured explicit formulas for Mayer and Ree-Hoover weights of the family $\left(K_n \setminus \left(\left(C_4 \cdot\cdot S_2\right)\cdot S_1\right)\right)$, $n \geq 7$, and more generally for the family $\left(K_n \setminus \left(\left(C_4 \cdot\cdot S_2\right)\cdot S_k\right)\right)$, $k \geq 1$, $n \geq k+6$. These formulas have been proved in Section 3 by the method of graph homomorphisms for the Ree-Hoover weight and by the linear relations between the two weights for Mayer's weight in Section 4. A similar work was done by the author, see [18], for families of graphs $\left(K_n \setminus \left(C_4 \cdot\cdot S_2\right)\right)$, $n \geq 7$ and $\left(K_n \setminus \left(S_1 \cdot C_4 \cdot\cdot S_2\right)\right)$, $n \geq 7$, and more generally for families $\left(K_n \setminus \left(S_k \cdot C_4 \cdot\cdot S_2\right)\right)$, $k \geq 1$, $n \geq k+6$. These developments pave the way for several future research prospects. For example, the extension of the exact calculation of Mayer's weight and Ree-Hoover's weight for families of graphs $\left(K_n \setminus \left(S_1 \cdot \left(C_4 \cdot\cdot S_2\right)\cdot S_1\right)\right)$, $n \geq 8$ and $\left(K_n \setminus \left(S_1 \cdot \left(C_4 \cdot\cdot S_2\right)\cdot S_2\right)\right)$, $n \geq 9$ and $\left(K_n \setminus \left(S_2 \cdot \left(C_4 \cdot\cdot S_2\right)\cdot S_1\right)\right)$, $n \geq 9$ and more generally for families $\left(K_n \setminus \left(S_1 \cdot \left(C_4 \cdot\cdot S_2\right)\cdot S_k\right)\right)$, $k \geq 1$, $n \geq k+6$ and $\left(K_n \setminus \left(S_k \cdot \left(C_4 \cdot\cdot S_2\right)\cdot S_1\right)\right)$, $k \geq 1$, $n \geq k+6$, with $\left(S_j \cdot \left(C_4 \cdot\cdot S_2\right)\cdot S_k\right)$ denote the graph obtained by joining with an edge of the graph $C_4 \cdot\cdot S_2$ the centers of a j-star and k-star.

Conflicts of Interest

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

References

[1] Kaouche, A. (2010) Invariants de graphes liés aux gaz imparfaits. Publications du Laboratoire de Combinatoire et d'Informatique Mathématiquee (LaCIM), Vol. 42.

[2] Labelle, J. (1981) Théorie des graphes. Modulo Éditeur.

[3] Leroux, P. (2004) Enumerative Problems Inspired by Mayer's Theory of Cluster Integrals. *The Electronic Journal of Combinatorics*, **11**, R32.

[4] Kaouche, A. and Leroux, P. (2008) Graph Weights Arising from Mayer and Ree-Hoover Theories. 20*th Annual International Conference on Formal Power Series and Algebraic Combinatorics*, 259-270.

[5] Kaouche, A. and Leroux, P. (2009) Mayer and Ree-Hoover Weights of Infinite Families of 2-Connected Graphs. *Séminaire Lotharingien de Combinatoire*, **61A**, B61Af.

[6] Labelle, G., Leroux, P. and Ducharme, M.G. (2007) Graph Weights Arising from

Mayer's Theory of Cluster Integrals. *Séminaire Lotharingien de Combinatoire*, **54**, B54m.

[7] Uhlenbeck, G.E. and Ford, G.W. (1963) Lectures in Statistical Mechanics. American Mathematical Society, Providence, 181 p.

[8] Ree, F.H. and Hoover, W.G. (1964) Fifth and Sixth Virial Coefficients for Hard Spheres and Hard Discs. *The Journal of Chemical Physics*, **40**, 939-950. https://doi.org/10.1063/1.1725286

[9] Ree, F.H. and Hoover, W.G. (1964) Reformulation of the Virial Series for Classical Fluids. *The Journal of Chemical Physics*, **41**, 1635-1645. https://doi.org/10.1063/1.1726136

[10] Ree, F.H. and Hoover, W.G. (1967) Seventh Virial Coefficients for Hard Spheres and Hard Discs. *The Journal of Chemical Physics*, **46**, 4181-4196. https://doi.org/10.1063/1.1840521

[11] Clisby, N. and McCoy, B.M. (2004) Negative Virial Coefficients and the Dominance of Loose Packed Diagrams for D-Dimensional Hard Spheres. *Journal of Statistical Physics*, **114**, 1361-1392. https://doi.org/10.1023/B:JOSS.0000013960.83555.7d

[12] Clisby, N. and McCoy, B.M. (2006) Ninth and Tenth Order Virial Coefficients for Hard Spheres in D Dimensions. *Journal of Statistical Physics*, **122**, 15-57. https://doi.org/10.1007/s10955-005-8080-0

[13] Clisby, N. (2004) Negative Virial Coefficients for Hard Spheres. Ph.D. Thesis, Stony Brook University, Stony Brook.

[14] Kaouche, A. (2016) Valeurs des poids de Mayer et des poids de Ree-Hoover pour tous les graphes 2-connexes de taille au plus 7 et leurs parametres descriptifs; Kaouche, A. (2016) Valeurs des poids de Mayer et des poids de Ree-Hoover pour tous les graphes 2-connexes de taille 8 et leurs parametres descriptifs. http://professeure.umoncton.ca/umce-kaouche_amel/files/umce-kaouche_amel/wf/wf/TableauRH7.pdf http://professeure.umoncton.ca/umce-kaouche_amel/files/umce-kaouche_amel/wf/wf/TableauRH8.pdf

[15] Kaouche, A. and Labelle, G. (2008) Mayer Polytopes and Divided Differences, Congrès Combinatorial Identities and Their Applications in Statistical Mechanics. Isaac Newton Institute de l'Université de Cambridge, en Angleterre. http://www.newton.ac.uk/webseminars/pg+ws/2008/csm/csmw03

[16] Kaouche, A. and Labelle, G. (2013) Mayer and Ree-Hoover Weights, Graphs Invariants and Bipartite Complete Graphs. *Pure Mathematics and Applications*, **24**, 19-29.

[17] Kaouche, A. and Labelle, G. (2014) Poids de Mayer et transformées de Fourier. *Annales Mathématiques du Québec Springer- Verlag*, **38**, 37-59. https://doi.org/10.1007/s40316-014-0018-y

[18] Kaouche, A. (2019) New Formulas for the Mayer and the Ree-Hoover Weights of Infinite Families of Graphs. *World Journal of Engineering and Technology*, **7**, 283-292. https://doi.org/10.4236/wjet.2019.72019

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1814-1825
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Quasi-Coordinate Search for a Randomly Moving Target

A. A. M. Teamah[1], W. A. Afifi[2,3*]

[1]Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt
[2]Mathematics and Statistics Department, College of Science, Taibah University, Yanbu, KSA
[3]Mathematics and Statistics Department, Canal Higher Institute of Engineering and Technology, Suez, Egypt
Email: *Walaa_affy@yahoo.com

How to cite this paper: Teamah, A.A.M. and Afifi, W.A. (2019) Quasi-Coordinate Search for a Randomly Moving Target. *Journal of Applied Mathematics and Physics*, **7**, 1814-1825.
https://doi.org/10.4236/jamp.2019.78124

Received: July 8, 2019
Accepted: August 16, 2019
Published: August 19, 2019

Abstract

In this paper, we study the quasi-coordinated search technique for a lost target assumed to move randomly on one of two disjoint lines according to a random walk motion, where there are two searchers beginning their search from the origin on the first line and other two searchers begin their search from the origin on the second line. But the motion of the two searchers on the first line is independent from the motion of the other two searchers on the second line. Here we introduce a model of search plan and investigate the expected value of the first meeting time between one of the searchers and the lost target. Also, we prove the existence of a search plan which minimizes the expected value of the first meeting time between one of the searchers and the target.

Keywords

Random Walker, Linear Search, Expected Value, Optimal Search Plane, Stochastic Process

1. Introduction

The searching for a lost target either located or moved is often a time-critical issue, that is, when the target is very important. The primary objective is to find and search for the lost target as soon as possible. The searching for lost targets has recently applications such as the search for a goldmine underground, the search for Landmines and navy mines, the search for the cancer cells in the human body, the search for missing black box of a plane crash in the depth of the sea of ocean, the search for a damaged unit in a large linear system such as telephone lines, and mining system, and so on [1] [2] [3]. Search problem when the

lost target is located or moved on the real line has been considered in [4]-[9]. The coordinated search technique discussed on the real line when the located target has symmetric or unsymmetric distribution as in [10] [11] [12]. Also, the coordinated search for a located target in the plane has been examined in [13] [14] [15] [16]. Recently, [17] and [18] proposed and studied a modern search model in the three-dimensional space to find a 3-D randomly located target by one searcher, two searchers and four searchers.

2. Problem Formulations

One of the most complicate problems when a mother loses her son in a way of multiple ways, here the primary objective is finding the lost son, as soon as possible in a minimum time. The survival rate of the son in this region gradually decreases, so the search team must organize itself quickly to begin the mission of the searching for the lost son immediately. Also, when the target is serious as a car, which filled by explosives, and it moves on one road from disjoint roads, and then the search effort must be unrestricted and we can use more than searcher to detect the target at right time.

The search team which consists of 4 searchers will organize itself on 2 straight lines to find the lost target as soon as possible. We clarify a modern technique by collaboration between each two searchers to find the lost person in minimum time. This problem can be characterized as follows.

2.1. The Searching Framework

The space of search: 2 disjoint lines.

The target: The target moves with a random walk motion on one of 2 disjoint straight lines.

The means of search: Looking for the lost target performed by two searches on each line. The searchers start searching for the target from the origins of the two lines with continuous paths and with equal speeds. In addition, the search spaces (2 straight lines) are separated into many distances.

2.2. The Searching Technique

Assume that we have two searchers S_1 and S_2 that start together looking for the lost target from O_1 on L_1. The two searchers coordinate their search about the lost target, where the searcher S_1 searches to the right and goes from the O_1 to H_1, and the searcher S_2 searches to the left and goes from O_1 to $-H_1$, the two searchers S_1 and S_2 reach to H_1 and $-H_1$ in the same time of G_1. Then they come back to O_1 again in the same time of G_2. If one of the two searchers do not find the lost target, then the two searchers S_1 and S_2 begin the new cycle search for the lost target, where they go from O_1 to H_2 and $-H_2$, respectively and they will reach to H_2 and $-H_2$ in the same time of G_3. Then they come back to O_1 again in the same time of G_4 and so on. Also, we have two other searchers S_3 and S_4 start together looking for the lost target from O_2 on the second line L_2, the searcher S_3

searchers to the right and goes from O_2 to $\bar{H}_1$, and the searchers S_4 searches to the left and goes to the left and goes from O_2 to $-\bar{H}_1$, the two searchers S_3 and S_4 reach to $\bar{H}_1$ and $-\bar{H}_1$ in the same time of $\bar{G}_1$. Then they come back to O_2 again in the same time of $\bar{G}_2$. If one of the two searchers not find the lost target, then the two searchers S_3 and S_4 begin the new cycle search for the lost target, where they go from O_2 to $\bar{H}_2$ and $-\bar{H}_2$, respectively and they will reach to $\bar{H}_2$ and $-\bar{H}_2$ in the same time of $\bar{G}_3$, then they come back to O_2 again in the same time of $\bar{G}_4$, and so on. The four searchers return to the O_1 and O_2 after searching successively common distances until the target is found.

2.3. The Movement of the Target and the Searchers

A target is assumed to move randomly on one of two disjoint lines according to a stochastic process $\{S(t), t \in I^+\}, I^+ = \{0,1,2,\cdots\}$. Assume that $\{Z_i\}_{i\geq 0}$ is a sequence of independent identically distributed random variables such as for any $i \geq 1$: $p(Z_i = 1) = p$ and $p(Z_i = -1) = 1 - p = q$, where $p, q > 0$. For $t > 0$, $t \in I^+$,

$$S(t) = \sum_{i=1}^{t} Z_i, S(0) = 0.$$

We assume the searchers S_1 and S_2 begin their search path from O_1 on L_1 with speeds V_1, and the searchers S_3 and S_4 begin their search path from O_2 on L_2 with speeds V_2, following the search paths which are functions $\phi_1 : R^+ \to R$ and $\bar{\phi}_1 : R^+ \to R$ on L_1 and $\phi_2 : R^+ \to R$ and $\bar{\phi}_2 : R^+ \to R$ on L_2, respectively, such that:

$$\left|\phi_1(t_1) - \phi_1(t_2)\right| = \left|\bar{\phi}_1(t_1) - \bar{\phi}_1(t_2)\right| \leq V_1 \left|t_1 - t_2\right|, \tag{1}$$

and

$$\left|\phi_2(t_1) - \phi_2(t_2)\right| = \left|\bar{\phi}_2(t_1) - \bar{\phi}_2(t_2)\right| \leq V_2 \left|t_1 - t_2\right|, \ \forall t_1, t_2 \in I^+, \tag{2}$$

where V_1 and V_2 are constants in R^+ and $\phi_1(0) = \bar{\phi}_1(0) = \phi_2(0) = \bar{\phi}_2(0) = 0$. Let the set of all search paths of the two searchers S_1 and S_2, which satisfy condition (1), be respectively by Φ_{v_1} and $\bar{\Phi}_{v_1}$ respectively and the set of all search paths of the searchers S_3 and S_4 which satisfy condition (2), be represented by Φ_{v_2} and $\bar{\Phi}_{v_2}$, respectively. we represented to the path of S_1 and S_2 by $\phi_0 = (\phi_1, \bar{\phi}_1) \in \Phi_0$ where $\bar{\phi}_0 = (\phi_2, \bar{\phi}_2) \in \bar{\Phi}_0$, where

$$\bar{\Phi}_0 = \left\{(\phi_2, \bar{\phi}_2) : \phi_2 \in \Phi_{v_2}, \bar{\phi}_2 \in \bar{\Phi}_{v_2}\right\}.$$

The search plan of the four searchers be represented by $\hat{\phi} = (\phi_0, \bar{\phi}_0) \in \hat{\Phi}$, where $\hat{\Phi} = \left\{(\phi_0, \bar{\phi}_0) : \phi_0 \in \Phi_0, \bar{\phi}_0 \in \bar{\Phi}_0\right\}$ is the set of all search plan.

We assume that $Z_0 = X$ if the target moves on L_1 and $Z_0 = Y$ if the target moves on L_2 such that $P(Z_0 = X) + P(Z_0 = Y) = 1$. There is a known probability measure $v_1 + v_2 = 1$ on $L_1 \bigcup L_2$ which describes the location of the target, where v_1 is probability measure induced by the position of the target on L_1, while v_2 on L_2. The first meeting time valued in I^+ defined as

$$\tau_{\hat{\phi}} = \inf\left\{t : \phi_1(t) = X + S(t) \text{ or } \bar{\phi}_1(t) = X + S(t) \right.$$
$$\left. \text{or } \phi_2(t) = Y + S(t) \text{ or } \bar{\phi}_2(t) = Y + S(t)\right\},$$

where Z_0 is a random variable representing the initial position of the target and valued in $2I$ (or $2I+1$) and independent of $S(t), t > 0$.

At the beginning of the search suppose that the lost target is existing on any integer point on L_1 but more than H_1 or less than $-H_1$ or the lost target is existing on an integer point on L_2 but more than $\bar{H}_1$ or less than $-\bar{H}_1$. Let τ_{ϕ_1} be the first meeting time between S_1 and the target and $\tau_{\hat{\phi}_1}$ be the first meeting time between S_2 and the target and τ_{ϕ_2} be the first meeting time between S_3 and the target and $\tau_{\hat{\phi}_2}$ be the first meeting time between S_4 and the target. The main objective is to find the search plan $\hat{\phi} = (\phi_0, \bar{\phi}_0) \in \hat{\Phi}$ such that $E(\tau_{\hat{\phi}}) < \infty$. In this case $\hat{\phi}$ is said to be a finite search plan, and if $E(\tau^*_{\hat{\phi}}) < E(\tau_{\hat{\phi}}), \forall \hat{\phi} \in \hat{\Phi}$, where E terms to expectation value, then we call $\hat{\phi}^*$ is an optimal search plan.

Given $n > 0$, if z is: $0 \le k_1 \le \dfrac{n+z}{2} \le n$, where k_1 is integer, then

$$p(S(n) = k_1) = \begin{cases} \binom{n}{k_1} p^{k_1} q^{n-k_1} \\ 0, \text{ if } k_1 \text{ does not exist} \end{cases}$$

2.4. Finite Search Plan

Let $\lambda_1, \lambda_2, \zeta_1, \zeta_2$ be positive integers such that $\zeta_1, \zeta_2 > 1$, $\lambda_1 = k\theta_1$, $\lambda_2 = k\theta_2$, where $k = 1, 2, \cdots$ and θ_1, θ_2 are the least positive integers and $V_1 = V_2 = 1$.

We shall define the sequences $\{G_i\}_{i\ge 0}, \{H_i\}_{i\ge 0}$ for the searcher S_1 on the first line L_1 and $\{\bar{G}_i\}_{i\ge 0}, \{\bar{H}_i\}_{i\ge 0}$ for the searcher S_3 on the second line L_2 and the search plans with speeds 1 as follows:

$$G_i = 2^{\frac{1}{2}\left[1-(-1)^{i-1}\right]} \lambda_1 \left(\zeta_1^{\frac{i}{2}+\frac{1}{4}-(-1)^i\frac{1}{4}} - 1 \right), \quad H_i = G_{2i-1}, \; i \ge 1 \text{ on } L_1,$$

$$\bar{G}_i = 2^{\frac{1}{2}\left[1-(-1)^{i-1}\right]} \lambda_2 \left(\zeta_2^{\frac{i}{2}+\frac{1}{4}-(-1)^i\frac{1}{4}} - 1 \right), \quad \bar{H}_i = \bar{G}_{2i-1}, \; i \ge 1 \text{ on } L_2.$$

We shall define the search path as follows:

for any $t \in I^+$, if $G_i \le t < G_{i+1}$, then

$$\phi_1(t) = \left(\frac{1}{2} H_{\frac{i+1}{2}}\right) + (-1)^{i+1} \left(\frac{1}{2} H_{\frac{i+1}{2}}\right) + (-1)^i (t - G_i),$$

and

$$\bar{\phi}_1(t) = -\phi_1(t).$$

Also, if $\bar{G}_i \le t < \bar{G}_{i+1}$, then

$$\phi_2(t) = \left(\frac{1}{2} \bar{H}_{\frac{i+1}{2}}\right) + (-1)^{i+1} \left(\frac{1}{2} \bar{H}_{\frac{i+1}{2}}\right) + (-1)^i (t - \bar{G}_i),$$

and

$$\bar{\phi}_2(t) = -\phi_2(t).$$

We define the notion

$\varphi_1(t)=S(t)-t,\ \tilde{\varphi}_1(t)=S(t)+t$ on L_1,

$\varphi_2(t)=S(t)-t,\ \tilde{\varphi}_2(t)=S(t)+t$ on L_2,

the searchers S_1 and S_2 return to the origin of L_1 after searching successively common distances $H_1,H_2,H_3,\cdots$, and $-H_1,-H_2,-H_3,\cdots$, respectively and the searchers S_3 and S_4 return to the origin of L_2 after searching successively common distances $\bar{H}_1,\bar{H}_2,\bar{H}_3,\cdots$, and $-\bar{H}_1,-\bar{H}_2,-\bar{H}_3,\cdots$, respectively until the target is found.

Theorem 1: If $\hat{\phi}=(\phi_0,\bar{\phi}_0)\in\hat{\Phi}$ is a search plan defined above, then the expectation $E(\tau_{\hat{\phi}})$ if finite if

$$w_1(x)=\sum_{i=1}^{\infty}\left(\zeta_1^i-1\right)p\left(\tilde{\varphi}_1(G_{2i-1})<-x\right),$$

$$w_2(x)=\sum_{i=1}^{\infty}\left(\zeta_1^i-1\right)p\left(\varphi_1(G_{2i-1})>-x\right),$$

$$w_3(x)=\sum_{i=1}^{\infty}\left(\zeta_1^i\left(\zeta_1^i-2\right)+1\right)p\left(\tilde{\varphi}_1(G_{2i})<-x\right),$$

$$w_4(y)=\sum_{i=1}^{\infty}\left(\zeta_1^i\left(\zeta_1^i-2\right)+1\right)p\left(\varphi_1(G_{2i})>-x\right),$$

$$w_5(y)=\sum_{i=1}^{\infty}\left(\zeta_2^i-1\right)p\left(\tilde{\varphi}_2(G_{2i-1})<-y\right),$$

$$w_6(y)=\sum_{i=1}^{\infty}\left(\zeta_2^i-1\right)p\left(\varphi_2(\tilde{G}_{2i-1})>-y\right),$$

$$w_7(y)=\sum_{i=1}^{\infty}\left(\zeta_2^i\left(\zeta_2^i-2\right)+1\right)p\left(\tilde{\varphi}_2(\tilde{G}_{2i})<-y\right),$$

and

$$w_8(y)=\sum_{i=1}^{\infty}\left(\zeta_2^i\left(\zeta_2^i-2\right)+1\right)p\left(\varphi_2(\tilde{G}_{2i})>-y\right). \tag{3}$$

are finite.

Proof: Assume that X and Y are independent of $S(t),t>0$, if $X>0$, then $X+S(t)>\phi_1(t)$ until the first meeting between S_1 and the target on L_1, also if $X<0$, then $X+S(t)<\hat{\phi}_1(t)$ until the first meeting between S_2 and the target on L_2. We can apply this assumption on the second line by replacing X by Y and $\phi_1,\bar{\phi}_1$ by $\phi_2,\bar{\phi}_2$ respectively. Hence, for any $i\geq 0$

$$p\left(\tau_{\hat{\phi}}>t\right)=p\left(\tau_{\phi_0}>t \text{ or } \tau_{\hat{\phi}_0}>t\right),$$

hence

$$\begin{aligned}E\left(\tau_{\hat{\phi}}\right)&=\int_0^{\infty}p\left(\tau_{\hat{\phi}}>t\right)\mathrm{d}t\\&\leq\sum_{i=0}^{\infty}\left[\int_{G_i}^{G_{i+1}}p\left(\tau_{\phi_0}>G_i\right)\mathrm{d}t+\int_{\tilde{G}_i}^{\tilde{G}_{i+1}}p\left(\tau_{\hat{\phi}_0}>\tilde{G}_i\right)\mathrm{d}t\right]\\&=\sum_{i=0}^{\infty}\left(2^{\frac{1}{2}\left[1-(-1)^{i+2}\right]}\lambda_1\left(\zeta_1^{\frac{i+1}{2}+\frac{1}{4}-(-1)^{i+1}\frac{1}{4}}-1\right)\right.\end{aligned}$$

$$
\begin{aligned}
&-2^{\frac{1}{2}\left[1-(-1)^{i+1}\right]}\lambda_1\left(\zeta_1^{\frac{i}{2}+\frac{1}{4}-(-1)^i\frac{1}{4}}-1\right)\Bigg)p\left(\tau_{\hat{\phi}_0}>G_i\right)\\
&+\Bigg(2^{\frac{1}{2}\left[1-(-1)^{i+2}\right]}\lambda_2\left(\zeta_2^{\frac{i+1}{2}+\frac{1}{4}-(-1)^{i+1}\frac{1}{4}}-1\right)\\
&-2^{\frac{1}{2}\left[1-(-1)^{i+1}\right]}\lambda_2\left(\zeta_2^{\frac{i}{2}+\frac{1}{4}-(-1)^i\frac{1}{4}}-1\right)\Bigg)p\left(\tau_{\hat{\phi}_0}>\tilde{G}_i\right)\\
&=\lambda_1\Big[\left(\left(\zeta_1-2\right)+1\right)p\left(\tau_{\phi_0}>0\right)+\left(\zeta_1-1\right)p\left(\tau_{\phi_0}>G_1\right)\\
&+\left(\zeta_1\left(\zeta_1-2\right)+1\right)p\left(\tau_{\phi_0}>G_2\right)+\left(\zeta_1^2-1\right)p\left(\tau_{\phi_0}>G_3\right)\\
&+\left(\zeta_1^2\left(\zeta_1-2\right)+1\right)p\left(\tau_{\phi_0}>G_4\right)+\left(\zeta_1^3-1\right)p\left(\tau_{\phi_0}>G_5\right)\\
&+\left(\zeta_1^3\left(\zeta_1-2\right)+1\right)p\left(\tau_{\phi_0}>G_6\right)+\cdots\Big]\\
&+\lambda_2\Big[\left(\left(\zeta_2-2\right)+1\right)p\left(\tau_{\bar{\phi}_0}>0\right)+\left(\zeta_2-1\right)p\left(\tau_{\bar{\phi}_0}>\bar{G}_1\right)\\
&+\left(\zeta_2\left(\zeta_2-2\right)+1\right)p\left(\tau_{\bar{\phi}_0}>\bar{G}_2\right)+\left(\zeta_2^2-1\right)p\left(\tau_{\bar{\phi}_0}>\bar{G}_3\right)\\
&+\left(\zeta_2^2\left(\zeta_2-2\right)+1\right)p\left(\tau_{\bar{\phi}_0}>\bar{G}_4\right)+\left(\zeta_2^3-1\right)p\left(\tau_{\bar{\phi}_0}>\bar{G}_5\right)\\
&+\left(\zeta_2^3\left(\zeta_2-2\right)+1\right)p\left(\tau_{\bar{\phi}_0}>\bar{G}_6\right)+\cdots\Big]
\end{aligned}
\tag{4}
$$

to solve Equation (4) we shall find the value of $p\left(\tau_{\phi_0}>G_{2i-1}\right)$, $p\left(\tau_{\bar{\phi}_0}>\bar{G}_{2i-1}\right)$, $p\left(\tau_{\phi_0}>G_{2i}\right)$ and the value of $p\left(\tau_{\bar{\phi}_0}>\bar{G}_{2i}\right)$ as the following

$$
\begin{aligned}
p\left(\tau_{\phi_0}>G_{2i-1}\right)\le&\int_{-\infty}^{0}p\left(x+S\left(G_{2i-1}\right)<-H_i\,/\,X=x\right)v_1\left(\mathrm{d}x\right)\\
&+\int_{0}^{\infty}p\left(x+S\left(G_{2i-1}\right)>H_i\,/\,X=x\right)v_1\left(\mathrm{d}x\right)
\end{aligned}
$$

We get

$$
\begin{aligned}
p\left(\tau_{\phi_0}>G_{2i-1}\right)\le&\int_{-\infty}^{0}p\left(\tilde{\varphi}_1\left(G_{2i-1}\right)<-x\right)v_1\left(\mathrm{d}x\right)\\
&+\int_{0}^{\infty}p\left(\varphi_1\left(G_{2i-1}\right)>-x\right)v_1\left(\mathrm{d}x\right)
\end{aligned}
\tag{5}
$$

also,

$$
\begin{aligned}
p\left(\tau_{\bar{\phi}_0}>\bar{G}_{2i-1}\right)\le&\int_{-\infty}^{0}p\left(\tilde{\varphi}_2\left(\bar{G}_{2i-1}\right)<-y\right)v_2\left(\mathrm{d}y\right)\\
&+\int_{0}^{\infty}p\left(\varphi_2\left(\bar{G}_{2i-1}\right)>-y\right)v_2\left(\mathrm{d}y\right)
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
p\left(\tau_{\phi_0}>G_{2i}\right)\le&\int_{-\infty}^{0}p\left(X+S\left(G_{2i}\right)<-2H_i\right)v_1\left(\mathrm{d}x\right)\\
&+\int_{0}^{\infty}p\left(x+S\left(G_{2i}\right)>2H_i\right)v_1\left(\mathrm{d}x\right)
\end{aligned}
\tag{7}
$$

We get

$$
p\left(\tau_{\phi_0}>G_{2i}\right)\le\int_{-\infty}^{0}p\left(\tilde{\varphi}_1\left(G_{2i}\right)<-x\right)v_1\left(\mathrm{d}x\right)+\int_{0}^{\infty}p\left(\varphi_1\left(G_{2i}\right)>-x\right)v_1\left(\mathrm{d}x\right)
\tag{8}
$$

$$p\left(\tau_{\phi_0} > \bar{G}_{2i}\right) \le \int_{-\infty}^{0} p\left(\tilde{\varphi}_1\left(\bar{G}_{2i}\right) < -y\right) v_2(\mathrm{d}y) + \int_{0}^{\infty} p\left(\varphi_1\left(\bar{G}_{2i}\right) > -y\right) v_2(\mathrm{d}y) \qquad (9)$$

substituting by (5), (6), (7) and (8) in (4) we can get

$$\begin{aligned}
E\left(\tau_{\hat{\phi}}\right) \le \lambda_1 \Big[& \left((\zeta_1 - 2) + 1\right) p\left(\tau_{\phi_0} > 0\right) + \left(\zeta_1 - 1\right) p\left(\tau_{\phi_0} > G_1\right) \\
& + \left(\zeta_1\left(\zeta_1 - 2\right) + 1\right) p\left(\tau_{\phi_0} > G_2\right) + \left(\zeta_1^2 - 1\right) p\left(\tau_{\phi_0} > G_3\right) \\
& + \left(\zeta_1^2\left(\zeta_1 - 2\right) + 1\right) p\left(\tau_{\phi_0} > G_4\right) + \left(\zeta_1^3 - 1\right) p\left(\tau_{\phi_0} > G_5\right) \\
& + \left(\zeta_1^3\left(\zeta_1 - 2\right) + 1\right) p\left(\tau_{\phi_0} > G_6\right) + \cdots \Big] \\
+ \lambda_2 \Big[& \left((\zeta_2 - 2) + 1\right) p\left(\tau_{\bar{\phi}_0} > 0\right) + \left(\zeta_2 - 1\right) p\left(\tau_{\bar{\phi}_0} > \bar{G}_1\right) \\
& + \left(\zeta_2\left(\zeta_2 - 2\right) + 1\right) p\left(\tau_{\bar{\phi}_0} > \bar{G}_2\right) + \left(\zeta_2^2 - 1\right) p\left(\tau_{\bar{\phi}_0} > \bar{G}_3\right) \\
& + \left(\zeta_2^2\left(\zeta_2 - 2\right) + 1\right) p\left(\tau_{\bar{\phi}_0} > \bar{G}_4\right) + \left(\zeta_2^3 - 1\right) p\left(\tau_{\bar{\phi}_0} > \bar{G}_5\right) \\
& + \left(\zeta_2^3\left(\zeta_2 - 2\right) + 1\right) p\left(\tau_{\bar{\phi}_0} > \bar{G}_6\right) + \cdots \Big]
\end{aligned}$$

hence

$$\begin{aligned}
E\left(\tau_{\hat{\phi}}\right) \le \lambda_1 \Bigg[& \left((\zeta_1 - 2) + 1\right) p\left(\tau_{\phi_0} > 0\right) + \left\{ \int_{-\infty}^{0} w_1(x) v_1(\mathrm{d}x) \right. \\
& \left. + \int_{0}^{\infty} w_2(x) v_1(\mathrm{d}x) \right\} + \left\{ \int_{-\infty}^{0} w_3(y) v_2(\mathrm{d}y) + \int_{0}^{\infty} w_4(y) v_2(\mathrm{d}y) \right\} \Bigg] \\
+ \lambda_2 \Bigg[& \left((\zeta_2 - 2) + 1\right) p\left(\tau_{\bar{\phi}_0} > 0\right) + \left\{ \int_{-\infty}^{0} w_5(y) v_2(\mathrm{d}y) \right. \\
& \left. + \int_{0}^{\infty} w_6(y) v_2(\mathrm{d}y) \right\} + \left\{ \int_{-\infty}^{0} w_7(y) v_2(\mathrm{d}y) + \int_{0}^{\infty} w_8(y) v_2(\mathrm{d}y) \right\} \Bigg]
\end{aligned}$$

where,

$$w_1(x) = \sum_{i=1}^{\infty} \left(\zeta_1^i - 1\right) p\left(\tilde{\varphi}_1\left(G_{2i-1}\right) < -x\right),$$

$$w_2(x) = \sum_{i=1}^{\infty} \left(\zeta_1^i - 1\right) p\left(\varphi_1\left(G_{2i-1}\right) > -x\right),$$

$$w_3(x) = \sum_{i=1}^{\infty} \left(\zeta_1^i\left(\zeta_1 - 2\right) + 1\right) p\left(\tilde{\varphi}_1\left(G_{2i}\right) < -x\right),$$

$$w_4(x) = \sum_{i=1}^{\infty} \left(\zeta_1^i\left(\zeta_1 - 2\right) + 1\right) p\left(\varphi_1\left(G_{2i}\right) > -x\right),$$

$$w_5(y) = \sum_{i=1}^{\infty} \left(\zeta_2^i - 1\right) p\left(\tilde{\varphi}_2\left(\bar{G}_{2i-1}\right) < -y\right),$$

$$w_6(y) = \sum_{i=1}^{\infty} \left(\zeta_2^i - 1\right) p\left(\varphi_2\left(\bar{G}_{2i-1}\right) > -y\right),$$

$$w_7(y) = \sum_{i=1}^{\infty} \left(\zeta_2^i\left(\zeta_2 - 2\right) + 1\right) p\left(\tilde{\varphi}_2\left(\bar{G}_{2i}\right) < -y\right),$$

and

$$w_8(y)=\sum_{i=1}^{\infty}\left(\zeta_2^i(\zeta_2-2)+1\right)p\left(\varphi_2\left(\bar{G}_{2i}\right)>-y\right).$$

Lemma 1: For any $k\geq 0$, let $a_n\geq 0$ for $n\geq 0$, and $a_{n+1}\leq a_n$. Let $\{d_n\}_{n\geq 0}$ be a strictly increasing sequence of integers with $d_0=0$,

$$\sum_{n=k}^{\infty}(d_{n+1}-d_n)a_{d_{n+1}}\leq\sum_{k=d_k}^{\infty}a_k\leq\sum_{n=k}^{\infty}(d_{n+1}-d_n)a_{d_n},$$

For more details see [1].

Theorem 2: The chosen search plan satisfies

$$w_1(x)\leq w_9(|x|),\ w_2(x)\leq w_{10}(|x|),$$

$$w_2(x)\leq w_{11}(|x|),\ w_4(x)\leq w_{12}(|x|),$$

$$w_5(y)\leq w_{13}(|y|),\ w_6(y)\leq w_{14}(|y|),$$

$$w_7(y)\leq w_{15}(|y|),\text{ and }w_8(y)\leq w_{16}(|y|),$$

where, $w_9(|x|)$, $w_{10}(|x|)$, $w_{11}(|x|)$, $w_{12}(|x|)$, $w_{13}(|y|)$, $w_{14}(|y|)$, $w_{15}(|y|)$, and $w_{16}(|y|)$ are linear function.

Proof: This theorem will prove for $w_2(x)$ and $w_6(y)$, and by similar way we can prove the other cases

$$w_2(x)=\sum_{i=0}^{\infty}\left(\zeta_1^i-1\right)p\left(\varphi_1\left(G_{2i-1}\right)>-x\right)$$

and

$$w_6(y)=\sum_{i=0}^{\infty}\left(\zeta_2^i-1\right)p\left(\varphi_2\left(\bar{G}_{2i-1}\right)>-y\right)$$

1) if $x\leq 0$, then

$$w_2(x)\leq w_2(0)$$

and if $y\leq 0$, then

$$w_6(y)\leq w_6(0),$$

2) if $x>0$, then

$$w_2(x)=w_2(0)+\sum_{i=0}^{\infty}\left(\zeta_1^i-1\right)p\left(-x<\varphi_1\left(G_{2i-1}\right)\leq 0\right),$$

and if $y>0$, then

$$w_6(y)=w_6(0)+\sum_{i=0}^{\infty}\left(\zeta_2^i-1\right)p\left(-y<\varphi_2\left(\bar{G}_{2i-1}\right)\leq 0\right),$$

from Theorem (2), see (Mohamed [1]) we obtain

$$w_2(0)=\sum_{i=0}^{\infty}\left(\zeta_1^i-1\right)p\left(\varphi_1\left(G_{2i-1}\right)>0\right)\leq\sum_{i=1}^{\infty}\left(\zeta_1^i-1\right)\varepsilon^{G_{2i-1}},\ 0<\varepsilon<1$$

and

$$w_6(0)=\sum_{i=0}^{\infty}\left(\zeta_2^i-1\right)p\left(\varphi_2\left(\bar{G}_{2i-1}\right)>0\right)\leq\sum_{i=1}^{\infty}\left(\zeta_2^i-1\right)\varepsilon^{\bar{G}_{2i-1}},\ 0<\varepsilon<1$$

Let us define the following

1) $V(n)=\varphi_1(n\theta_1)/2=\sum_{i=1}^{n}W_i$, where $\{W_i\}$ is a sequence of (i. i. d. r. v.) $\bar{V}(n)=\varphi_2(n\theta_2)/2=\sum_{i=1}^{n}\bar{W}_i$, where $\{\bar{W}_i\}$ is a sequence of (i. i. d. r. v.).

2) $d_n=G_{2n-1}/\theta_1=k(\zeta_1^n-1)$, $\bar{d}_n=\bar{G}_{2n-1}/\theta_2=k(\zeta_2^n-1)$.

3) $a(n)=\frac{n}{n+k}p(-x/2<V(n)\le 0)=\sum_{i=0}^{|x|/2}p[-(j+1)<V(n)\le(-j)]$,

$\bar{a}(n)=\frac{n}{n+k}p(-y/2<\bar{V}(n)\le 0)=\sum_{i=0}^{|y|/2}p[-(j+1)<\bar{V}(n)\le(-j)]$,

4) m_1 is an integer such that $dm_1=b_1|x|+b_2$, and m_2 is an integer such that $dm_2=\bar{b}_1|y|+\bar{b}_2$,

5) $\alpha_1=\frac{\zeta_1}{(\zeta_1-1)k}$, and $\alpha_2=\frac{\zeta_2}{(\zeta_2-1)k}$,

and

6) $U_1(j,j+1)=\sum_{n=0}^{\infty}p[-(j+1)<V(n)<(-j)]$,

$\bar{U}_1(j,j+1)=\sum_{n=0}^{\infty}p[-(j+1)<\bar{V}(n)\le(-j)]$,

then $U_1(j,j+1)$ and $\bar{U}_1(j,j+1)$ satisfies the condition of the renewal equation, for more details see [19].

If $n>d_{m_1}$ and $n>d_{m_2}$ then by Theorem (2) see (Mohamed [1]) $a(n)$ and $\bar{a}(n)$ are non increasing and we can apply Lemma (2) to obtain

$$\begin{aligned}
w_2(x)-w_2(0)&=\sum_{i=1}^{\infty}(\zeta_1^i-1)p(-x<\varphi_1(G_{2i-1})\le 0)\\
&=\sum_{n=1}^{n_1}\zeta_1^n a(d_n)+\sum_{n=n_1+1}^{\infty}\zeta_1^n a(d_n)\\
&\le\sum_{n=1}^{n_1}\zeta_1^n+\alpha_1\sum_{n=n_1+1}^{\infty}(d_n-d_{n-1})a(d_n)\\
&\le\sum_{n=1}^{n_1}\zeta_1^n+\alpha_1\sum_{n=d_{m_1}}^{\infty}a(n)\\
&\le\sum_{n=1}^{n_1}\zeta_1^n+\alpha_1\sum_{n=d_{m_1}}^{\infty}\sum_{i=0}^{|x|/2}p[-(j+1)<V(n)\le(-j)]\\
&\le\sum_{n=1}^{n_1}\zeta_1^n+\alpha_1\sum_{j=0}^{|x|/2}U_1(j,j+1)
\end{aligned}$$

and

$$\begin{aligned}
w_6(x)-w_6(0)&=\sum_{i=0}^{\infty}(\zeta_2^i-1)p(-y<\varphi_2(G_{2i-1})\le 0)\\
&=\sum_{n=1}^{n_2}\zeta_2^n a(\bar{d}_n)+\sum_{n=n_2+1}^{\infty}\zeta_2^n a(\bar{d}_n)\\
&\le\sum_{n=1}^{n_2}\zeta_2^n+\alpha_2\sum_{n=n_2+1}^{\infty}(\bar{d}_n-\bar{d}_{n-1})a(\bar{d}_n)
\end{aligned}$$

$$\leq \sum_{n=1}^{n_2} \zeta_2^n + \alpha_2 \sum_{n=d_{m_2}}^{\infty} \bar{a}(n)$$

$$\leq \sum_{n=1}^{n_2} \zeta_2^n + \alpha_2 \sum_{j=0}^{|y|/2} \bar{U}_1(j, j+1)$$

Since $U_1(j, j+1)$ and $\bar{U}_1(j, j+1)$ satisfied the condition of the renewal equation, hence $U_1(j, j+1)$ and $\bar{U}_1(j, j+1)$ is bounded for all j by a constant, so

$$w_2(x) \leq w_2(0) + N_1 + N_2|x| = w_{10}(|x|),$$

and

$$w_6(x) \leq w_6(0) + \bar{N}_1 + \bar{N}_2|x| = w_{14}(|y|).$$

Theorem 3: If $\hat{\phi} = (\phi_0, \bar{\phi}_0) \in \hat{\Phi}$ is a finite search plan, then $E|Z_0|$ is finite.

Proof: If $E(\tau_{\hat{\phi}}) < \infty$, then $p(\tau_{\hat{\phi}} \text{ is finite}) = 1$ and so

$$p(\tau_{\phi_0} \text{ is finite}) + p(\tau_{\bar{\phi}_0} \text{ is finite}) = 1,$$

then, we conclude that

$$p(\tau_{\phi_0} \text{ is finite}) = 1 \quad \text{and} \quad p(\tau_{\bar{\phi}_0} \text{ is finite}) = 0,$$

or

$$p(\tau_{\phi_0} \text{ is finite}) = 0 \quad \text{and} \quad p(\tau_{\bar{\phi}_0} \text{ is finite}) = 1.$$

On the first line L_1 if $p(\tau_{\phi_0} \text{ is finite}) = 1$, then $X_0 = \phi(\tau_{\phi_0}) - S(\tau_{\phi_0})$ with probability one and hence

$$E|X_0| \leq E(\tau_{\phi_0}) + E|S(\tau_{\phi_0})|.$$

If $E(\tau_{\phi_0}) < \infty$, but $|S(\tau_{\phi_0})| \leq \tau_{\phi_0}$, then $E|S(\tau_{\phi_0})| \leq E(\tau_{\phi_0})$ and $E|X_0| < \infty$.

On the second line L_2 if $p(\tau_{\bar{\phi}_0} \text{ is finite}) = 1$, then $Y_0 = \phi(\tau_{\bar{\phi}_0}) - S(\tau_{\bar{\phi}_0})$ with probability one, by the same way we can get $E|Y_0|$ is finite on the second line L_2.

3. Existence of an Optimal Search Plan

Theorem 4: Let for any $t \in I^+$, let $S(t)$ be a process. The mapping $\hat{\phi} \to E(\tau_{\hat{\phi}}) \in R^+$ is lower semi-continuous on $\hat{\Phi}(t)$.

Proof: Let $I(\hat{\phi}, t)$ be the indicator function of the set $\{\tau_{\hat{\phi}} \geq t\}$ by the Fatou Lebesque theorem see (Stone [16]) we get

$$E(\tau_{\hat{\phi}}) = E\left[\sum_{t=1}^{\infty} I(\hat{\phi}, t)\right]$$

$$= E\left[\sum_{t=1}^{\infty} \liminf_{i \to \infty} I(\hat{\phi}_n, t)\right] \leq \liminf_{i \to \infty} E(\tau_{\hat{\phi}}),$$

for any sequence $\hat{\phi}_n \to \hat{\phi}$ in $\hat{\Phi}(t)$ is sequentially compact [20], thus the mapping $\hat{\phi} \to E(\tau_{\hat{\phi}})$ is lower semi continuous on $\hat{\Phi}(t)$, then this mapping attains its minimum.

4. Conclusions

We have described a new kind of search technique to find a lost moving target on one of two disjoint lines. The motion of the four searchers on the two lines in the quasi-coordinated search technique is independent, and this helps us to find the lost target without waste of time and cost, especially if this target is valuable as the search for lost children. Actually we calculated the finite search plan. Also; we proved the existence of an optimal search plan which minimizes the expected value of the first meeting time between one of the searchers and the target.

In the future work, we will introduce an important search problem, looking for a randomly moving target as a general case and the searchers will begin their mission from any point on the line.

Conflicts of Interest

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

References

[1] Mohamed, A.A. and El-Rayes, A.B. (1989) Search for a Randomly Moving Target. *The Third ORMA Conference*, Vol. 2, 323-329.

[2] Alpern, S. and Howard, J.V. (2000) Alternating Search at Two Locations. *Dynamics and Control*, **10**, 319-339. https://doi.org/10.1023/A:1011245715521

[3] El-Hadidy, M.A. and El-Bagoury, A.H. (2016) Optimal Search Strategy for a Three-Dimensional Randomly Located Target. *International Journal of Operational Research*, **29**, 115-126. http://www.inderscience.com/link.php?id=83178 https://doi.org/10.1504/IJOR.2017.10003932

[4] Mohamed, A.A. (2005) The Generalized Search for One Dimensional Random Walker. *International Journal of Pure and Applied Mathematics*, **19**, 375-387.

[5] Mohamed, A.A. and Abou-Gabal, H.M. (2003) Linear Search with Multiple Searchers for a Randomly Moving Target. *International Conference for Statistics, Computer Science and its Application*, 115-124.

[6] Mohamed, A.A. and Abou-Gabal, H.M. (2004) Multiplicative Linear Search Problem. *Egyptian Statistical Journal, Cairo University*, **48**, 34-45.

[7] Beck, A. and Warren, P. (1973) The Return of the Linear Search Problem. *Israel Journal of Mathematics*, **14**, 169-183. https://doi.org/10.1007/BF02762672

[8] Balkhi, Z.T. (1989) The Generalized Optimal Search Paths for Continuous Univariate Random Variable. *Journal of the Operations Research*, **23**, 67-96. https://doi.org/10.1051/ro/1989230100671

[9] Stone, L.D. (1975) Theory of Optimal Search. Academic Press, New York.

[10] Mohamed, A.A., Abou-Gabal, H.M. and Afifi, W.A. (2013) Double Coordinate Search Problem. *International Journal of Contemporary Mathematical Science*, **8**.

[11] Mohamed, A.A., Abou-Gabal, H.M. and Afifi, W.A. (2016) Generalized Coordinated Search for a Randomly Located Target. *Delta Journal of Science*, **38**.

[12] Reyniers, D.J. (1996) Coordinated Search for an Object on the Line. *European Journal of Operational Research*, **95**, 663-670. https://doi.org/10.1016/S0377-2217(96)00314-1

[13] Mohamed, A.A. and El-Hadidy, M. (2013) Coordinated Search for a Conditionally Deterministic Target Motion in the Plan. *European Journal of Mathematical Sciences*, **2**, 272-295.

[14] Mohamed, A.A., Abou-Gabal, H.M. and El-Hadidy, M. (2009) Coordinated Search for a Randomly Located Target on the Plane. *European Journal of Pure and Applied Mathematics*, **2**, 97-111.

[15] Mohamed, A.A., Fergany, H.A. and El-Hadidy, M. (2012) On the Coordinated Search Problem on the Plane. *Istanbul University Journal of the School of Business Administration*, **41**, 80-102.

[16] Bourgault, F., Furukawa, T. and Durrant-Whyte, H. (2003) Coordinated Decentralized Search for a Lost Target in a Bayesian World. *Proceedings IEEERSJ International Conference*, *Intelligent Robots and Systems*, Vol. 1.

[17] Mohamed, A.A. and EL-Bagoury, A.H. (2019) Minimizing the Expected Time to Detect a Randomly Located Lost Target Using 3-Dimensional Search Technique. *Journal of Communications in Statistics*.

[18] Mohamed, A.A., El-Hadidy, M. and EL-Bagoury, A.H. (2017) 3-Dimensional Coordinated Search Technique for a Randomly Located Target. *International Journal of Computing Science and Mathematics*, **9**.
https://doi.org/10.1504/IJCSM.2018.093152

[19] Feller, W. (1966) An Introduction to Probability Theory and Its Applications. Second Edition, Wiley, New York.

[20] Mohamed, A.A., El-Rayes, A.B. and Abou-Gabal, H.M. (2003) Linear Search for a Brownian Target Motion. *Acta Mathematica Scientia*, **23B**, 321-327.
https://doi.org/10.1016/S0252-9602(17)30338-7

Journal of Applied Mathematics and Physics, 2019, 7, 1826-1839
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Method of Construction of Material That Work on All the Range of Wavelengths or Frequency or Energy of Photon

Ram Chandra Pageni, Saddam Husain Dhobi, Narayan Panthi, Sonam Gyaljen Tamang, Sudarshan Shrestha

Department of Physics, Tribhuvan University, Kathmandu, Nepal
Email: pagenirc09@gmail.com, saddamdhobe@gmail.com, narayan.panthi321@gmail.com, gyaljen511@gmail.com, sudson292017@gmail.com

How to cite this paper: Pageni, R.C., Dhobi, S.H., Panthi, N., Tamang, S.G. and Shrestha, S. (2019) Method of Construction of Material That Work on All the Range of Wavelengths or Frequency or Energy of Photon. *Journal of Applied Mathematics and Physics*, **7**, 1826-1839.
https://doi.org/10.4236/jamp.2019.78125

Received: July 4, 2019
Accepted: August 16, 2019
Published: August 19, 2019

Abstract

The main objective of this research work is to decrease work function of any given element or compound or material. To decrease the work function of the given material we have to decrease the bandwidth between conduct band and valance band. Because according to definition of work function, the amount of energy that required the remove the electron from valance band of an atom and it is also called ionization energy. These all energies depend upon the band width that is greater than the band width greater energy required to remove the electron from the surface, and less than the band width and lesser amount of energy required to remove the electron from of materials. In this work we are trying to give an theoretical model or relation, how to decrease the work function of a material by applying external pressure on atoms and doping of the material that has screening or shielding effects. With the help of this model we can increase the efficiency of material used in solar cell that is cell work for all range of frequencies and by construction material bases on this we can increase the efficiency of solar cell or any type of material working solar cell principle.

Keywords

Conduction Band, Ionization Energy, Solar Cell, Valance Band, Work Function

1. Introduction

The total amount of energy received at ground level from the Sun is about 3.3% higher than average in January and lower in July. In terms of energy, sunlight at

Earth's surface is around 52 to 55 percent infrared 42 to 43 percent visible, and 3 to 5 percent ultraviolet. In this research work, we are trying to give a theoretical model, that decreases the work function of the material especially, photonic material from IR to Visible range frequencies because the maximum amount of energy in our atmosphere is energy of photon. The utilization of such huge amount of energy is possible only if we constructed materials which can emit the electron from atom, for all ranges of frequencies. This is possible to decrease the work function of the material either by doping of the different material that has shielding effect or by applying the Van Der Waals force relation. Base on these principles we can construct photonic materials that have different work function and can be used for different frequency of photon especially IR and Visible photon.

Photonic crystals are periodically structured electromagnetic media, generally possessing photonic band gaps: ranges of frequency in which light cannot propagate through the structure. The study of photonic crystals is likewise governed by the Bloch F. theorem, and intentionally introduced defects in the crystal. Felix Bloch pioneered the study of wave propagation in three dimension-ally periodic media in 1928, which proved that waves in such a medium could propagate without scattering, when the photon incident on the material the energy of the photon observed electron or electron cloud of an atom. The energy observation is started from the outer most valance orbital *i.e.* the outer most electron of an atom and goes on the excited state or kicks out from the orbits of an atom; hence the photoelectric effect take place. Let the energy of incident photon on the surface of photonic material is hf, where h is plank constant and f is frequency and the amount of energy required to remove or electron from its valance orbit to conduct orbit hf_1 where f_1 is frequency needed to excite the electron from valance band or orbit1 and the energy hf_2 of excited state, where f_2 is frequency of exciting the electron for conducing band or orbit 2 and the difference between them gives, $hf = hf_2 - hf_1$, which implies smaller the difference of $hf_2 - hf_1$, where hf is smaller that is f is smaller and shows that on changing the energy level or orbit difference we can obtain the necessary value of $hf_2 - hf_1$. In this way we can change frequency for our desire and increase the efficiency of the photonic materials like solar cell.

On the other hand, work function of the material is define as the minimum amount of the energy required to remove the electron from the material surface which is given by hf_3 where h = plank's factor and f_3 is the minimum frequency required to remove the electron from the surface, which is equal to or less than incident frequency; then we can obtain the relation.

$$hf_2 - hf_1 = hf_3 \le hf \tag{1}$$

In general,

$$hf_{n+1} - hf_n = hf_{n'} \le hf \tag{2}$$

In Equation (2), f_{n+1} is frequency of orbit $n+1$ and f_n is the frequency of orbit

n.

Now on putting the value of $f = \frac{c}{\lambda}$ where c = velocity of light and λ in Equation (2) we obtain the relation

$$1/\lambda_{n+1} - 1/\lambda_n = 1/\lambda_{n'} = 1/\lambda \tag{3}$$

where λ_{n+1}, λ_n and λ_n, or λ is corresponding wavelength of $n + 1$, n orbit and difference wavelength of two orbit $n + 1$ and n.

Photonic crystals usually consist of dielectric materials, that is, materials that serve as electrical insulators or in which an electromagnetic field can be propagated with low loss. Materials used for making a Photonic band gap are: Silicon, Germanium, Gallium and so on.

2. Literature

Composite optical materials can display useful optical properties that are qualitatively dissimilar from those of their underlying constituents. Nano composite materials are especially well suited for photonic applications because they can be constructed in such a manner as to produce enhanced nonlinear optical response. More recently entangled states of light have been used to perform functions unthinkable in the context of classical physics, such as the demonstration of quantum cryptography and quantum teleportation (those materials which interact with certain frequency that is we can transform quantum information from one to another place). One application entails the development of techniques for the construction of imaging systems that can achieve a transverse resolution that exceeds the classical Rayleigh criterion [1]. Photonic crystal working in the optical range of electromagnetic spectrum but PCs for X-ray should present a modulation of some A^0 which is a solid state crystal. EM field store more energy if it has extreme at the region of high dielectric function. Bragg diffraction is the primary feature of the system of PBG material can be used for initial characterization. When refractive index contrast is low it is good approx. to use ordinary diffraction or dynamic to characterized PCs [2]. The mid-infrared wavelength range from 2 to 20 microns is a spectral region of tremendous interest which is important for a wide range of applications ranging from chemical and bio-Sensing to spectroscopy and thermal imaging, optical fiber can transmit the light from visible to near IR, industrial and military applications, such as remote sensing and explosive detection or free-space communication systems [3].

The electromagnetic band gap overlaps the electronic band edge by at least a few kT in energy, then electron-hole radioactive recombination, which creates periodic three-dimensional dielectric structures in which there exists an electromagnetic band gap. The recombination rate of electrons and holes in a semiconductor can be expressed as a power series in the injected carrier density. The combined electron and hole three-body Auger recombination coefficient are intrinsic to the material. In today's high quality, double hetero structure has low

enough defect density [4]. The concept of a photonic material is explained in terms of Bragg diffraction. The concept of Bragg reflection applies equally well to visible radiation except that we cannot rely on atoms to do work for all researcher [5]. GaAs are used in various electronic applications such as wireless communication, microwave, and high-speed digital systems. Silicon appears as thin films of amorphous or poly crystalline form in TFT-technology, in solar cell. SiGe is needed to introduce electron accelerating strain in the Si lattice. GaAs, InP, and GaP, has occurred culminating in the development of a blue semiconductor laser using GaN [6]. As in the 3D case, we can create dielectric system which exhibits a complete photonic band gap, a set of frequencies in which light of any polarization cannot propagate in any in-plane direction and drill air columns in dielectric. At the dielectric contrast of GaAs, the only combination which was found to have a photonic band gap in both polarization was the triangular lattice of air columns in dielectric [7]. Photonic crystals can be used to selectively reflect specific frequencies of electromagnetic radiation. It has recently been shown experimentally that geometrically asymmetric silicon structures support strong near-infrared magnetic resonances akin to the familiar 'trapped mode' of metallic asymmetric split ring designs. Silicon and silicon nit-ride are assumed lossless with refractive indices of 3.5 and 2.0 respectively in the near-infrared range under consideration. All in-plane optical forces generated within the meta-material structure are canceled and only out-of plane forces act on the dielectric beams. These drive each beam to move up or down until elastic restoring forces balance the optical forces. Optical absorption and variations in ambient temperature may lead to thermo-mechanical changes in the structure [8] [9].

Far-infrared radiation traveling through a polar crystal, such as GaAs, couples strongly to optical phonon. The transverse optical (TO) and longitudinal optical (LO) phonon define a region of negative permittivity and the center of a region of high loss, the so-called "Reststrahlen band" of the material. Optical transparency in the 20 - 60 μm range for intrinsic III-V semiconductors is limited by absorption due to optical phonon. Consequently, materials and devices based on III-V semiconductors are scarce in the Reststrahlen region [10]. Beyond semiconductors, superconductors also have the potential to advance the state-of the-art in far-IR detection. Microwave kinetic inductance detectors sense incident radiation by detecting small changes in the inductance of a thin superconductor from the breaking of Cooper pairs. The small changes in the inductance can be detected by monitoring the resonance frequency of a microwave resonator that incorporates the superconducting film. A challenge associated with Microwave kinetic inductance detectors for the far-IR is coupling incident radiation into the device [11].

A well-known thin-film Gallium nit-ride LED structure is fabricated by removing the sapphire layer using a laser and roughening the revealed n-doped Gallium nit-ride layer. Gallium nit-ride thin-films, wafer-bonded to oxidized

silicon wafers, provide a promising platform for nonlinear integrated optics. The obtainable quality factors in optical micro-resonators based on bonded Gallium nit-ride thin films are limited by the remaining surface roughness. Aluminum nit-ride, on the other hand, is a wide-band gap semiconductor with the capability of direct integration on silicon (100) substrates with suitable mechanical and thermal properties [12] [13].

Photonic crystals have a periodic dielectric structure with high index contrast, designed to control photons in the same way that conventional crystals control electrons. PCs also possess a band gap, a so-called photonic band gap, where the material acts like a photonic insulator and light of certain frequencies cannot propagate. The Electro optics effect is the change in the index of refraction of a material with an applied external electric field, given by $\delta\text{-}\, n_{ij} = -n_{ij}^2\left(r_{ijk}E_k + s_{ijkl}E_kE_l\right)$ [14]. Transparent electrodes are used to apply an electric field, which, in combination with polarized, may be used to either block or transmit the light. Electro optic crystals have extremely small Electro optic coefficients. Liquid crystal optical-phased array beam steers tend to be slow, provide non-continuous diffraction steering, and have a very limited steering range because thick LC layers are problematic. Acoustic optic beam steers have a larger steering range but are also diffraction, require very large power supplies and expensive crystals [15].

Photonic crystals are meta-materials designed to display a periodic modulation of the refractive index. One of the intrinsic shortcomings of photonic crystals is the highly selective reflection from Bragg planes due to crystalline symmetries. For example, dye-free reflective color displays, colored packing materials or cosmetics are preferentially non-iridescent and thus non-crystalline. For a photonic crystal of thickness L the sum of transmittance T and spectacular reflectance R is $T + R = 1$. The gap wavelength λ_G and the transmittance displays an exponential decay $T = \mathrm{e}^{-L/L_B}$, where L_B is the Bragg length [16].

AlGaN alloys have been the default choice for the development of Deep ultraviolet Op-to-electronic devices. Deep ultraviolet light emitting diodes LED and laser diodes with high quantum efficiencies QEs is the low conductivity of p-type AlGaN. The resistivity of Mg-doped AlGaN increases with Al-content and becomes extremely high in Mg-doped AlN. Seebeck effect measurement is a well-established technique to distinguish between n-type and p-type conductivity of a semiconductor. The temperature gradient creates a voltage between the cold and hot ends due to the diffusion of thermally excited charged carriers [17].

As the investigation of photo current generation mechanisms in 2D opt o electronics goes deeper, 2D material-based devices have also been demonstrated to exhibit elevated performance. Over the past decade, it has been shown that most graphene based photo-transistors exhibit photo responsive around 10 mA/W [18]. First reported single-layer MoS_2 photo transistors reach responsibility to 7.5 mA/W with 50 V gate bias, and exhibit stable response time within 50 ms. Because mono layer MoS_2 is a direct-band gap semiconductor due to

quantum-mechanical confinement, it could be suitable for applications in opt o electronic devices where the direct band gap would allow a high absorption coefficient and efficient electron-hole pair generation under photo excitation. The maximum external photo is responsive of 880 A W-1 at a wavelength of 561 nm and a photo response in the 400 - 680 nm range. MoS_2 shows important potential for applications in MoS_2-based integrated opt o electronic circuits, light sensing, biomedical imaging, video recording and spectroscopy [19] [20].

Near the gap in a photonic band, gap material the effective index of refraction can become less than unity and in fact can approach zero at the band edge itself-leading to ultra-refractive optical effects. Ultra-refractive optics with photonic band materials has many applications, including laser accelerators and lenses of ultra-short focal lengths. Scully and colleagues, In current prototypes of PBG materials the structure is about 85% air and so losses are less than in a homogeneous dielectric and discussed the novel properties of the effective index of refraction of PBG material when operating near a band edge, using a quantitative and simple one-dimensional model of a three-dimensional PPG structure [21].

We present designs of 2D, isotropic, disordered, photonic materials of arbitrary size with complete band gaps blocking all directions and polarization. The designs with the largest band gaps are obtained by a constrained optimization method that starts from a hyper uniform disordered point pattern Since their introduction in 1987, photonic band gap (PBG) materials have evolved dramatically, and their unusual properties have led to diverse applications. Obtaining complete PBGs in dielectric materials without long-range order is counter intuitive. If the arrangement of dielectrics has local geometric order, a tight binding model with nearly uniform coefficients describes the propagation of light in the limit of high dielectric constant ratio. Weaire and Thorpe proved that band gaps could exist in continuous random tetrahedral coordinated networks, commonly used as models for amorphous silicon and germanium. The comparison to electronic band gaps is also useful in comparing states near the band edges and continuum. For a perfectly ordered crystal (or photonic crystal), the electronic (photonic) states at the band edge are propagating such that the electrons (electromagnetic fields) sample many sites. If modest disorder is introduced, localized states begin to fill in the gap so that the states just below and just above are localized. Although formally the disordered hetero-structures do not have equivalent propagating states, an analogous phenomenon occurs [22].

The radiation effects on Ge-doped and (Fluorine) F-doped fibers and performs: the first play a crucial role in the photo sensitivity property, the second improves the dielectric radiation hardness even at low concentrations. Usually classical optical fibers for telecommunications are used in the IR, from 835 to 1600 nm, medical application and plasma diagnostic use of visible and UV region were optical transmission is affected by many losses. Fluorine is a promising key material in optical fiber technology directed to applications requiring

high and stable transmission in UV and visible spectral range [23]. Photons interact very weakly with transparent optical media. Due to their fast propagation speed, photons use for transmitting quantum information and distributing quantum correlations. Additionally, photons possess many degrees of freedom; include spatial and temporal modes, frequency, polarization and angular momentum. Photons also possess continuous quantum variables, such as the quantized field quadrature, which can be utilized in special classes of information protocols. Silicon has a much higher refractive index than silica, leading to 1000 times smaller wave guide bend radius compared to that in silica wave guides. Silicon's indirect band gap of 1.12 eV and low intrinsic carrier concentration makes it transparent to photons at the telecommunication wavelength (1.55 μm) [24].

Up to now, entangled photon pairs have been generated by optical pumping in passive semiconductor wave guides by exploiting four-wave mixing in silicon or spontaneous parametric down conversion in aluminum gallium arsenide. Aluminum Gallium Arsenide device that emits photon pairs at telecom wavelengths and operates at room temperature [25]. Three-dimensional confinement of both electrons and holes in hetero-structures gives rise to quantum dots which potentially provide on-demand single-photon generation across the NIR range with near unity internal efficiency. Efficient photon collection represents a major technological challenge associated with quantum emitters. Integrated collection techniques with over 98% mode coupling efficiency exist for quantum dots [26].

InGaAs/GaAs quantum dots demonstrate single-photon emission for wavelengths up to 1400 nm, while InAsP/InP quantum dots emission can be achieved across the entire telecom spectrum. Multi-photon entangled states can be generated, constituting a particularly useful resource for quantum communication. The combination of a diode junction and an embedded quantum dot has led to the demonstration of an electrically driven entangled-photon source of high enough quality to perform quantum teleportation [27] [28].

It is also found that the size and composition with alloy effect can tune the band-gap energy, which suggests an effective way to reach the desirable electrical and optical properties. Interestingly, both the size and the composition can tune the band-gap energy of Nano compounds. In particular, the composition tun-able band-gap may be a better way for wider band-gap materials [29]. Ionization potential is the electric potential (V) required to separate an electron from the orbital system in free space with the kinetic energy remaining unchanged. Ionization energy is the work done in removing the electron at zero temperature and is measured conveniently in electron volts, where 1 eV = 1.6022×10^{19} J. Screening (electron-electron repulsion) reduces electron-nucleus attractions in helium and two-electron atomic ions but ionization energies are not functions of simple squares. First ionization energies vary in the order B > Al < Ga > In < Tl and C > Si > Ge > Sn < Pb but decrease with increasing atomic number down groups 15

to 18. In 1930, Slater give a relation $I = R\left(Z^*/N^*\right)^2$, where Z^* = an effective nuclear charge, N^* = an effective quantum number could be assigned to each electron and I = energy required to ionize an electron [30].

The photo-ionization of 3d" impurities in semiconductors with zinc blende structure is to be treated by investigating the matrix elements for electric dipole transitions between linear combinations of d-type one-electron wave functions which can be mixed with p-states and the continuum of states in the conduction band, *i.e.* Bloch functions [31]. The force attracting the electron to the nucleus depends upon magnitude of the core charge, and the separation of the electron from the nucleus. The first ionization energy will therefore be the work done when removing an electron from the core and so will be that needed to pull the electron away against the attraction from the core charge. This will not be an infinite quantity as the magnitude of the force falls rapidly with distance [32].

The ionization energy (IE) is defined as the minimum amount of energy that needs to be absorbed by an atom or molecule in its electronic and vibration ground states in order to form an ion that is also in its ground states by ejection of an electron. Ionization energies of most molecules are in the range of 7 - 15 eV. The ionization cross section describes an area through which the electron must travel in order to effectively interact with the neutral [33]. Lebedev gave the radiation pressure of light in 1901. In principle, if the currents in the mirror/antenna can be determined from the drive fields, then the electromagnetic fields due those currents can be calculated, and combined with the drive fields to obtain the total fields. This effect of the currents on themselves can be expressed mathematically as an integral equation for the currents, which takes into account the good-conductor boundary condition at the surface of the mirror/antenna [34].

3. Methodology: Theoretical Calculation

In this paper we are trying to give theoretical model or relation that help to construct such type of material which work on difference range of frequencies that is produce the current or energy for all frequencies. To give the theoretical relation here we follow the Bohr's relation of energy for hydrogen atom and pressure area relation.

If we increase the pressure on an atom surface externally, we can decrease the width or the distance between two outer most orbits of an atom that is the two outer most orbits come closure. This shows amount of energy is required to excite electron from inner outer most orbit to outer most orbit needed less. But in normal condition or without applying the pressure on atom the distance between the inner outer most and outer most is quite greater then applied presure condition and hence the energy required to excite the electron from inner most outer orbit to outer most orbit needed more.

Here the outer most orbit is last orbit of an atom when we consider the counting of orbit from nucleus and the inner outer most orbit is second last orbit of nucleus.

Mathematical Derivation:

From hydrogen model we have Bohr's energy for different orbits. Let us consider the energy of valance band or energy of ground state electron is E_n and conductance band or excited state electron is E_{n+1}.

Now we relation $E_{n+1} - E_n = hf$ when an electron goes from ground state to excited state and on coming to ground state again by emitting the energy of hf.

Similar phenomena we are using here that is amount of energy required to go electron from valance band to conduct band is given by $E_{n+1} - E_n = hf$, Where h = planks factor and f = incident frequency of a photon. Let us consider, N = Nucleus of an atom, n = 1 is the nearest orbit of nucleus where electron is bounded tightly, n = 2 is orbit of nucleus where electron is less bounded then n = 1, n = 3 is another orbit where electron is less bounded then n = 2. In similar way the electron bounded is going decrease as we go far away from nucleus. As shown in Figure 1.

According to Bohr's the amount of energy required to bounded the electron around the nucleus is given by

$$E_n = -\frac{1}{n^2}\frac{me^4z^2}{8\varepsilon^2h^2} = -\frac{z^2R_H}{n^2} \tag{4}$$

where R_H = Rydberg's constant.

We also have, Bohr's radius

$$r_n = \frac{n^2h^2}{z4\pi^2me^2} = 0.529\times\frac{n^2}{z}$$

This is the distance from the center of nucleus to the orbit of an electron revolving around the nucleus.

$$\text{Or,}\quad \frac{1}{n^2} = \frac{0.529}{zr_n} \tag{5}$$

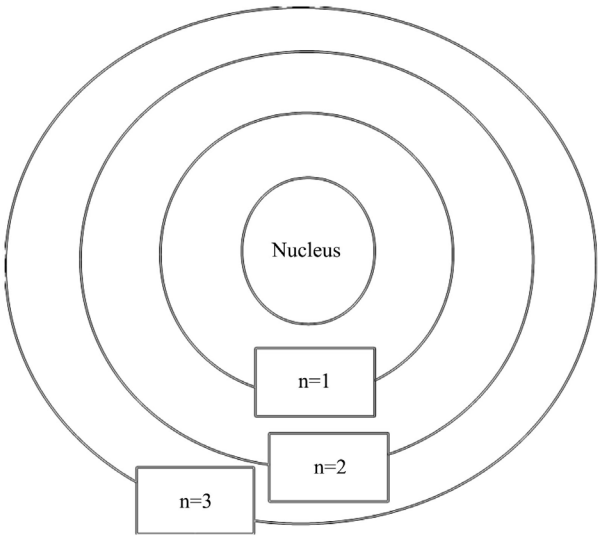


Figure 1. Energy Level for an atom without applying the external pressure.

Now From Equations (4) and (5) we have

$$E_n = -\frac{0.529zR_H}{r_n} \tag{6}$$

This relation shows that how much amount of energy required to bind the electron by proton or nucleus when the electron is at a certain *i.e.* r_n.

Let us consider, E_{n+1} is the amount of energy required to remove the electron from the outer shell of atom of any atom is at a distance r_{n+1} then the energy can be obtain similar as Equation (6).

Since we have an energy relation

$$E_{n+1} = -\frac{0.529zR_H}{r_{1+n}} \tag{7}$$

In other word, E_{n+1} and E_n is the energy level of valance band or ground band of orbit n and $n+1$ of an atom then the transition between them is possible. If we consider n as ground state and $n+1$ is excited state of electron in an atom then, Now on subtracting Equation (7) from (6) we have,

$$hf = E_{n+1} - E_n = 0.529zR_H\left(\frac{1}{r_n} - \frac{1}{r_{n+1}}\right) \tag{8}$$

where hf is incident photon, which is needed to, excited an electron from one energy level to another that means from E_n to E_{n+1} level.

$$hf = E_{n+1} - E_n = 0.529zR_H\left(\frac{r_{n+1} - r_n}{r_n \times r_{n+1}}\right) \tag{9}$$

Here from Equation (9) we have an interesting relation that the different between two energy levels are related to the distance between two orbits is directly proportional and product is inversely related which means incident photon energy is directly related difference of $r_{n+1} - r_n$ it means that smaller the difference smaller the amount of energy required to excited an electron from r_n to r_{n+1}. And another relation is inversely related to the product of the distance of two orbits from the nucleus ($r_n \square r_{n+1}$).

Since the atom is not solid that is within atom there is the space so we can compress the atom. Here we are studying the compression of the atom from the outer surface area during the compression the of an atom equal from all sides. the outer shell come to closure to the second last orbit or ground state orbit while the ground state orbit is less effect due to apply the pressure from outer side of atom.

Lets P_n and P_{n+1} are the pressure applying on energy level of an atom which is at a distance of r_n and r_{n+1} from nucleus, with constant force F then pressure exerted on n and $n+1$ orbits.

Then external pressure is applied on orbit of $n+1$ and n orbit with constant force to an atom is given by

$$P_n = \frac{F}{A_n} = -\frac{F}{4\pi r_n^2} \tag{10}$$

$$P_{n+1} = \frac{F}{A_{n+1}} = -\frac{F}{4\pi r_{n+1}^2} \tag{11}$$

respectively. Here negative sign in (10) and (11) implies show that force is acting from outside that mean in outer most orbit. Since the pressure is applied from outside or surface of an atom then pressure exerted on $n+1$ orbit is greater then pressure exerted on n orbit for constant force.

Now on subtracting Equation (10) from (11) we have,

$$P_{n+1} - P_n = \frac{F}{4\pi}\left(\frac{1}{r_n^2} - \frac{1}{r_{n+1}^2}\right) \tag{12}$$

$$P_{n+1} - P_n = \frac{F}{4\pi}\left(\frac{r_{n+1}^2 - r_n^2}{r_n^2 \times r_{n+1}^2}\right) \tag{13}$$

Now, from Equation (13) we have clearly, see that pressure different is directly related to different and inversely related to square product of distance from nucleus to electron.

$$P_{n+1} - P_n = \frac{F}{4\pi}\left(\frac{\left(r_{n+1} - r_n\right)\left(r_{n+1} + r_n\right)}{r_n^2 \times r_{n+1}^2}\right) \tag{14}$$

Now, on putting the value from Equation (9) on Equation (14), we have,

$$P_{n+1} - P_n = \frac{F}{4\pi}\left(\frac{\left(E_{n+1} - E_n\right)\left(r_{n+1} + r_n\right)}{0.529 z R_H \left(r_{n+1} \times r_n\right)}\right) \tag{15}$$

Since P_{n+1} is greater than P_n then $P_{n+1} - P_n$ must be positive which is also depend up on the $\left(E_{n+1} - E_n\right)$ of Equation (15) which is depend upon $(r_{n+1} - r_n)$.

It is possible because of the atom has space according to the Ruth-er forth scattering of an atom. Here we can press the electron for the certain radius because of repulsion of nuclear. In addition, Van Der Waals forces between atoms are depends upon the sized of atoms that is smaller the sized smaller the Van Der Waals force. Also on doping the material having the greater atomic number decrease attraction of the electron with the nucleus that is less amount of energy required to remove the electron from the surface of atom or materials. From the above Methodology, we can decrease the work function of the element or compound or materials.

4. Result and Conclusion

Hence, from the above relation (15), pressure, energy and orbit radius we can change the work function of material and make them work for all frequencies ranges. This helps to increase the efficiency of solar cell material or similar principle material. Once material is constructed or follows this principle, energy crisis should be ended forever. In other hand, we can also say that, same materials can be used for different range of frequencies and extracted energy from all ranges radiation. Hence such material can also solve the problem of energy where solar radiation is absent.

Acknowledgements

We would like to thanks all the members of Innovative Ghar Nepal and Robotic Academy Nepal who directly and indirectly provide us research space and support during our research work. Similarly, we would like to thanks our family member who provide peaceful environment during my research work.

Conflicts of Interest

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

References

[1] Boyd, R.W. (2003) Novel Photonic Materials for Advanced Imaging Applications. *Journal of the Korean Physical Society*, **43**, 603-605. https://www.researchgate.net/publication/228415943_Novel_Photonic_Materials_for_Advanced_Imaging_Applications

[2] Lopez, C. (2003) Material Aspect for Photoinc Crystal. *Advance Material*, **15**, 1679-1684. https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.200300386 https://doi.org/10.1002/adma.200300386

[3] Academy of Finland Research Programmes (2013) Foresight: Materials for Photonics. https://www.aka.fi/globalassets/32akatemiaohjelmat/oma/fotoniikka_ennakointiesite.pdf

[4] Yablonovitch, E. (1987) Inhibited Spontaneous Emission in Solid-State Physics and Electronics. *Physical Review Letters*, **58**, 2059-2061. https://www.iap.uni-jena.de/iapmedia/de/Lecture/Nanomaterials1367272800/NanoMat12_Paper_08.pdf https://doi.org/10.1103/PhysRevLett.58.2059

[5] Pendry, J.B. (1999) Photonic Gap Materials. *Current Science*, **76**, 1311-1136.

[6] Hoff, A.M. and Ruzyllo, J. (2006) Electronics and Photonics. In: *The Electrochemical Society Interface*, Springer, Berlin, 36-37.

[7] Meade, R.D., Brommer, K.D., Rappe, A.M., and Joannopoulos, J.D. (1992) Existence of a Photonic Band Gap in Two Dimensions. *Applied Physics Letter*, **61**, 495-496. https://aip.scitation.org/doi/10.1063/1.107868 https://doi.org/10.1063/1.107868

[8] Pouya, C., Overvelde, J.T.B., Kolle, M., Aizenberg, J., Bertoldi, K., Weaver, J.C. and Vukusic, P. (2015) Characterization of a Mechanically Tunable Gyroid Photonic Crystal Inspired by the Butterfly *Parides sesostris*. *Advanced Optical Materials*, **4**, 99-105. https://doi.org/10.1002/adom.201500436

[9] Zhang, J., MacDonald, K.F. and Zheludev, N.I. (2013) Nonlinear Dielectric Optomechanical Metamaterials. *Light*: *Science & Applications*, **2**, e96. https://doi.org/10.1038/lsa.2013.52

[10] Feng, K., Streyer, W., Zhong, Y., Hoffman, A.J. and Wasserman, D. (2015) Photonic Materials, Structures and Devices for Reststrahlen Optics. *Optics Express*, **23**, A1418. https://doi.org/10.1364/OE.23.0A1418

[11] Stevenson, T.R., Adams, J.S., Hsieh, W., Moseley, S.H., Travers, D.E., Uyen, K., Wollack, E.J. and Zmuidzinas, J. (2009) Superconducting Films for Absorber-Coupled

MKID Detectors for Sub-Millimeter and Far-Infrared Astronomy. *IEEE Transactions on Applied Superconductivity*, **19**, 561-564. https://doi.org/10.1109/TASC.2009.2019661

[12] Huang, H.W., Lin, C.H., Huang, Z.K., Lee, K.Y., Yu, C.C. and Kuo, H.C. (2010) Double Photonic Quasi-Crystal Structure Effect on GaN-Based Vertical-Injection Light-Emitting Diodes. *Japanese Journal of Applied Physics*, 49, Article ID: 022101. https://pdfs.semanticscholar.org/1b32/13ee213c10244da09e2cd1993270768a2bf6.pdf https://doi.org/10.1143/JJAP.49.022101

[13] Pernice, W.H.P., Xiong, C., Schuck, C. and Tang, H.X. (2012) High-q Aluminum Nitride Photonic Crystal Nanobeam Cavities. *Applied Physics Letter*s, **100**, Article ID: 091105. https://doi.org/10.1063/1.3690888

[14] Scrymgeour, D., Malkova, N., Kim, S. and Gopalan, V. (2003) Electro-Optic Control of the Superprism Effect in Photonic Crystals. *Applied Physics Letters*, **82**, 3176-3178. https://aip.scitation.org/doi/10.1063/1.1574402 https://doi.org/10.1063/1.1574402

[15] Davis, S.R., Rommel, S.D., Farca, G. and Anderson, M.H. (2008) A New Generation of Previously Unrealizable Photonic Devices as Enabled by a Unique Electro-Optic Waveguide Architecture. *SPIE*, Vol. 7050, 8-9.

[16] Muller, N., Haberko, J., Marichy, C. and Scheffold, F. (2013) Silicon Hyper Uniform Disordered Photonic Materials with a Pronounced Gap in the Shortwave Infrared. *Advanced Optical Materials*, **2**, 115-119. https://doi.org/10.1002/adom.201300415

[17] Dahal, R., Li, J., Majety, S., Pantha, B.N., Cao, X.K., Lin, J.Y. and Jiang, H.X. (2011) Epitaxially Grown Semiconducting Hexagonal Boron Nitride as a Deep Ultraviolet Photonic Material. *Applied Physics Letters*, **98**, Article ID: 211110. https://aip.scitation.org/doi/10.1063/1.3593958 https://doi.org/10.1063/1.3593958

[18] Xia, F., Mueller, T., Lin, Y.M., Garcia, A.V. and Avouris, P. (2009) Ultrafast Graphene Photodetector. *Nature Nanotechnology*, **4**, 839-843. http://www.gr-sci.net/papers/nnano.2009.292.pdf https://doi.org/10.1038/nnano.2009.292

[19] Yin, Z., Li, H., Li, H., Jiang, L., Shi, Y., Sun, Y., Lu, G., Zhang, Q., Chen, X. and Zhang, H. (2011) Single-Layer MoS_2 Phototransistors. *ACS Nano*, **6**, 74-80. https://pubs.acs.org/doi/10.1021/nn2024557 https://doi.org/10.1021/nn2024557

[20] Sanchez, O.L., Lembke, D., Kayci, M., Radenovic, A. and Kis, A. (2013) Ultrasensitive Photodetectors Based on Monolayer MoS_2. *Nature Nanotechonology*, **8**, 497-499. https://www.nature.com/articles/nnano.2013.100 https://doi.org/10.1038/nnano.2013.100

[21] Dowling, J.P. and Bowden, C.M. (1994) Anomalous Index of Refraction in Photonic Bandgap Materials. *Journal of Modern Optics*, **41**, 345-351. https://doi.org/10.1080/09500349414550371 https://www.tandfonline.com/doi/abs/10.1080/09500349414550371

[22] Florescu, M., Torquato, S. and Steinhardt, P.J. (2009) Designer Disordered Materials with Large. Complete Photonic Band Gaps. *Proceedings of the National Academy of Sciences Proceedings of the National Academy of Sciences*, **106**, 20658-20663. https://doi.org/10.1073/pnas.0907744106

[23] Origlio, G., Girard, S., Boscaino, R., Boukenter, A., Cannas, M. and Ouerdane, Y. (2009) Optical and Photonic Material Hardness for Energetic Environments. *UVX*, **2008**, 127-132. https://doi.org/10.1051/uvx/2009020

[24] Bogdanov, S., Shalaginov, M.Y., Boltasseva, A. and Shalaev, V.M. (2017) Material Platforms for Integrated Quantum Photonics. *Optical Materials Express*, **7**, 111-132. https://arxiv.org/pdf/1610.00729

[25] Boitier, F., Orieux, A., Autebert, C., Lemaître, A., Galopin, E., Manquest, C., Sirtori, C., Favero, I., Leo, G. and Ducci, S. (2014) Electrically Injected Photon-Pair Source at Room Temperature. *Physical Review Letters*, **112**, Article ID: 183901. https://link.aps.org/doi/10.1103/PhysRevLett.112.183901 https://doi.org/10.1103/PhysRevLett.112.183901

[26] Arcari, M., Söllner, I., Javadi, A., Hansen, S.L., Mahmoodian, S., Liu, J., Thyrrestrup, H., Lee, E.H., Song, J.D., Stobbe, S. and Lodahl, P. (2014) Near-Unity Coupling Efficiency of a Quantum Emitter to a Photonic Crystal Waveguide. *Physical Review Letters*, **113**, Article ID: 093603. https://link.aps.org/doi/10.1103/PhysRevLett.113.093603 https://doi.org/10.1103/PhysRevLett.113.093603
Salter, C.L., Stevenson, R.M., Farrer, I., Nicoll, C.A., Ritchie, D.A. and Shields, A.J. (2010) An Entangled-Light Emitting Diode. *Nature*, **465**, 594-597. https://www.researchgate.net/publication/44646833_An_entangled_light-emitting_diode https://doi.org/10.1038/nature09078

[27] Nilsson, J., Stevenson, R.M., Chan, K.H.A., Szymanska, J.S., Lucamarini, M., Ward, M.B., Bennett, A.J., Salter, C.L., Farrer, I., Ritchie, D.A. and Shields, A.J. (2013) Quantum Teleportation Using a Light Emitting Diode. *Nature Photonics*, **7**, 311-315. https://www.nature.com/articles/nphoton.2013.10?draft=journal https://doi.org/10.1038/nphoton.2013.10

[28] Wang, Y., Ouyang, G., Wang, L.L., Tang, L.M., Tang, D.S. and Sun, C.Q. (2008) Size and Composition Induced Band-Gap Change of Nano Structured Compound of II-VI Semiconductors. *Chemical Physics Letters*, **463**, 383-386. https://www3.ntu.edu.sg/home/ecqsun/RTF/CPL_WangY.pdf https://doi.org/10.1016/j.cplett.2008.08.083

[29] Lang, P.F. and Smith, B.C. (2003) Ionization Energies of Atoms and Atomic Ions. *Journal of Chemical Education*, **80**, 938-946. https://pubs.acs.org/doi/abs/10.1021/ed080p938 https://doi.org/10.1021/ed080p938

[30] Grebe, G. and Schulz, H.J. (2013) Luminescence of Cr^{2+} Centres and Related Optical Interactions Involving Crystal Field Levels of Chromium Ions in Zinc Sulfide. *Verlag Zeitschrift für Naturforschung*, **29**, 1808-1812.

[31] Taber, K.S. (2003) Understanding Ionisation Energy: Physical, Chemical and Alternative Conceptions. *Chemistry Education: Research and Practice*, **4**, 155-156. http://www.uoi.gr/cerp/2003_May/pdf/05Taber.pdf https://doi.org/10.1039/B3RP90010J

[32] Gross, J. (2011) Mass Spectrometry. 2nd Edition, Springer-Verlag, Berlin, Heidelberg, 26-29. https://doi.org/10.1007/978-3-642-10711-5

[33] McDonald, K.T. (2009) Radiation Pressure of a Monochromatic Plane Wave on a Flat Mirror. Joseph Henry Laboratories, Princeton University, Princeton, 1-6.

Journal of Applied Mathematics and Physics, 2019, 7, 1840-1860
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Dynamic Analysis for a SIQR Epidemic Model with Specific Nonlinear Incidence Rate

Jie Xu, Tiansi Zhang

College of Science, University of Shanghai for Science and Technology, Shanghai, China
Email: 819924208@qq.com

How to cite this paper: Xu, J. and Zhang, T.S. (2019) Dynamic Analysis for a SIQR Epidemic Model with Specific Nonlinear Incidence Rate. *Journal of Applied Mathematics and Physics*, **7**, 1840-1860.
https://doi.org/10.4236/jamp.2019.78126

Received: July 17, 2019
Accepted: August 20, 2019
Published: August 23, 2019

Abstract

The article investigates a SIQR epidemic model with specific nonlinear incidence rate and stochastic model based on the former, respectively. For deterministic model, we study the existence and stability of the equilibrium points by controlling threshold parameter R_0 which determines whether the disease disappears or prevails. Then by using Routh-Hurwitz criteria and constructing suitable Lyapunov function, we get that the disease-free equilibrium is globally asymptotically stable if $R_0 < 1$ or unstable if $R_0 > 1$. In addition, the endemic equilibrium point is globally asymptotically stable in certain region when $R_0 > 1$. For the corresponding stochastic model, the existence and uniqueness of the global positive solution are discussed and some sufficient conditions for the extinction of the disease and the persistence in the mean are established by defining its related stochastic threshold R_0^S. Moreover, our analytical results show that the introduction of random fluctuations can suppress disease outbreak. And numerical simulations are used to confirm the theoretical results.

Keywords

Epidemic Model, Specific Nonlinear Incidence Rate, Lyapunov Function, Stability, Existence, Persistence

1. Introduction

Infectious diseases have always been a thorny issue that endangers human health, triggers social unrest and even affects national stability. Therefore, it is of great significance to take effective prevention measures to control the epidemic by establishing mathematical models with typical characteristics, discovering the transmission and development trends of infectious diseases. In the last decades,

many authors have made a great headway on SIR (Susceptible-Infected-Removed) epidemic models [1] [2] [3]. However, for some diseases such as SARS, smallpox, foot-and-mouth disease, parrot fever and so on, introducing quarantine is one of the most pivotal and effective control means. In a model, assuming that some susceptible individuals become infected ones, and then the infected individuals flow into three parts, some remains at class I, some move into the removed R after recovering health, while some are transferred to the quarantine class Q and enter class R until they are no longer infectious. It is worth noting that the infected and quarantined people have permanent immunity after recovery. The model with the above characteristics is called SIQR (Susceptible-Infected-Quarantined-Removed) model. It is not difficult to find that a large amount of researches on the impact of quarantine on infectious diseases have been carried out so far [4] [5] [6] [7]. And in 2017, Joshi *et al.* [8] studied a SIQR epidemic model with saturated incidence rate and proved the global stability of the disease-free and endemic equilibrium.

In order to realistically reflect the process of human-to-human disease transmission, it is very important to determine the specific form of the incidence function which describes the increased number of infected people per unit time and plays a vital role in epidemiological dynamics research. Due to the complexity of disease transmission in real life, many scholars admit that the nonlinear incidence function is more reasonable than the bilinear incidence and standard incidence. The specific nonlinear incidence $\frac{\beta S(t)I(t)}{f(S(t),I(t))}$ was proposed in 2013 [9], where $f(S(t),I(t))=1+\alpha_1 S(t)+\alpha_2 I(t)+\alpha_3 S(t)I(t)$ and $\alpha_1,\alpha_2,\alpha_3$ are saturation factors that measure psychological or inhibitory effects and non-negative constants. Obviously, depending on the values of $\alpha_1,\alpha_2,\alpha_3$, the incidence can be changed to various common types of incidence rates in existing literatures, including the bilinear incidence rate, the saturation incidence, Beddington-DeAngelis incidence [10] and Crowley-Martin response [11]. Therefore, it is more interesting and valuable than the saturation incidence and is also widely used to study epidemic diseases. For instance, Adnani *et al.* [12] introduced the effect of white noise into a SIRS epidemic model with the above incidence. They analyzed the global existence, positivity and boundedness of solutions, as well as the dynamics of stochastic model. And Hattaf *et al.* [13] applied the incidence to a stochastic delayed SIR epidemic model with temporary immunity. They proved that the model is mathematically and biologically well-posed and also obtained sufficient conditions for the extinction and persistence of the disease. However, there are few articles on the SIQR model with this incidence rate.

In this paper, to improve and generalize the model of Joshi *et al.* [8], we propose a new SIQR epidemic model based on the incidence [9]. The deterministic differential equations of the model are as follows:

$$\begin{cases} \dfrac{\mathrm{d}S(t)}{\mathrm{d}t} = A - \mu S(t) - \dfrac{\beta S(t) I(t)}{f\left(S(t), I(t)\right)}, \\ \dfrac{\mathrm{d}I(t)}{\mathrm{d}t} = \dfrac{\beta S(t) I(t)}{f\left(S(t), I(t)\right)} - (\delta + \gamma + \mu + \mu_1) I(t), \\ \dfrac{\mathrm{d}Q(t)}{\mathrm{d}t} = \delta I(t) - (\rho + \mu + \mu_2) Q(t), \\ \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma I(t) + \rho Q(t) - \mu R(t). \end{cases} \tag{1}$$

The total population $N(t)$ is divided into four compartments and $N(t) = S(t) + I(t) + Q(t) + R(t)$, where $S(t), I(t), Q(t)$ and $R(t)$ are the number of the susceptible, infected, quarantined and recovered individuals at time t, respectively. The parameter constants have the following biological meanings: A is the recruitment rate of the susceptible through birth and immigration; μ is the natural death rate of the population; μ_1 is the disease-caused mortality of infective individuals; μ_2 is the disease-caused mortality of quarantined individuals; β represents contact rate of an infected person with other compartment members per unit time; δ is the isolation rate of the compartment I quarantined directly to enter Q; γ is the recovery rate of infected individuals; ρ is the recovery rate of quarantined individuals. In addition, all parameters of model (1) are supposed to be nonnegative constants. Especially, A and μ are positive constants.

In fact, any system is more or less affected by environmental factors. Stochastic models can predict the future dynamics of the system accurately compared to their corresponding deterministic models. Therefore, when establishing population model, many stochastic biological systems and stochastic epidemic models have been presented and studied [14] [15]. One of the most main ways to introduce random effects is to directly perturb the parameters of the deterministic model by Gaussian white noise. As an expansion of model (1), now we introduce white noises into (1) by substituting the parameters μ_i, β with $\mu_i + \sigma_i B_i(t)$ $(i = 1, 2)$ and $\beta + \sigma_3 B_3(t)$, where $B(t) = \left(B_1(t), B_2(t), B_3(t)\right)$ is a standard Brownian motion. $\sigma_i^2 > 0$ $(i = 1, 2, 3)$ denote the intensity of the white noise. Other parameters are the same as in model (1). Hence, the stochastic system is described by

$$\begin{cases} \mathrm{d}S(t) = \left(A - \mu S(t) - \dfrac{\beta S(t) I(t)}{f\left(S(t), I(t)\right)} \right) \mathrm{d}t - \dfrac{\sigma_3 S(t) I(t)}{f\left(S(t), I(t)\right)} \mathrm{d}B_3(t), \\ \mathrm{d}I(t) = \left(\dfrac{\beta S(t) I(t)}{f\left(S(t), I(t)\right)} - (\delta + \gamma + \mu + \mu_1) I(t) \right) \mathrm{d}t - \sigma_1 I(t) \mathrm{d}B_1(t) \\ \qquad + \dfrac{\sigma_3 S(t) I(t)}{f\left(S(t), I(t)\right)} \mathrm{d}B_3(t), \\ \mathrm{d}Q(t) = \left(\delta I(t) - (\rho + \mu + \mu_2) Q(t)\right) \mathrm{d}t - \sigma_2 Q(t) \mathrm{d}B_2(t), \\ \mathrm{d}R(t) = \left(\gamma I(t) + \rho Q(t) - \mu R(t)\right) \mathrm{d}t. \end{cases} \tag{2}$$

This paper is organized as follows. In Section 2, we present some preliminaries which will be used in our following analysis. In Section 3, the existence and stability of the equilibrium points of deterministic system is analyzed. In Section 4, we study dynamics of the stochastic model. Firstly, the existence and uniqueness of the global positive solution is proved. Then, the extinction and persistence of the disease under certain conditions is discussed. Finally, numerical simulations are presented to illustrate our main results. Section 5 just provides a brief discussion and the summary.

2. Preliminaries

In this section, some notations, definitions and lemmas are provided to prove our main results. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (*i.e.* it is increasing and right continuous while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets). And $B_i(t)(i=1,2,3)$ are defined on this complete probability space.

Consider the 4-dimensional stochastic differential equation

$$\mathrm{d}x(t) = f(x(t),t)\mathrm{d}t + g(x(t),t)\mathrm{d}B(t) \quad \text{for } t \geq t_0, \tag{3}$$

with initial value $x_0 \in \mathbb{R}_+^4$. We define the differential operator L of Equation (3) as follows:

$$L = \frac{\partial}{\partial t} + \sum_{i=1}^{4} f_i(x,t)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{4}\left[g^{\mathrm{T}}(x,t)g(x,t)\right]_{ij}\frac{\partial^2}{\partial x_i \partial x_j}.$$

Let L act on a nonnegative function $V(x,t) \in \mathcal{C}^{2,1}\left(\mathbb{R}_+^4 \times [t_0,\infty); \mathbb{R}_+\right)$. Then,

$$LV(x,t) = V_t(x,t) + V_x(x,t)f(x,t) + \frac{1}{2}trace\left[g^{\mathrm{T}}(x,t)V_{xx}(x,t)g(x,t)\right],$$

where $\mathbb{R}_+^4 = \{x_i > 0, i = 1,2,3,4\}$. By Itô's formula, $\mathrm{d}V(x,t) = LV(x,t)\mathrm{d}t + V_x(x,t)g(x,t)\mathrm{d}B(t)$. For an integrable function χ on $[0,+\infty)$, we define

$$\langle \chi(t) \rangle = \frac{1}{t}\int_0^t \chi(s)\mathrm{d}s.$$

Definition 2.1 *System* (2) *is said to be persistent in the mean if*

$$\liminf_{t\to\infty} \langle I(t) \rangle > 0 \ \ a.s..$$

Moreover, we need the following lemma (see Lemma 5.1 in [16]).

Lemma 2.2. *Let* $g \in \mathcal{C}(\mathbb{R}_+ \times \Omega, \mathbb{R})$ *and* $G \in \mathcal{C}(\mathbb{R}_+ \times \Omega, \mathbb{R})$. *If there exist two real numbers* $\lambda_0 \geq 0$ *and* $\lambda > 0$ *for all* $t \geq 0$, *such that*

$$\ln g(t) \geq \lambda_0 t - \lambda\int_0^t g(s)\mathrm{d}s + G(t) \ \text{ and } \ \lim_{t\to\infty}\frac{G(t)}{t} = 0 \ \ a.s.,$$

then

$$\liminf_{t\to\infty} \langle g(t) \rangle \geq \frac{\lambda_0}{\lambda} \ \ a.s..$$

Lemma 2.3. *Consider the following two systems*

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t,x),\ \frac{\mathrm{d}y}{\mathrm{d}t} = g(y),$$

where $x, y \in \mathbb{R}^n$, f and g are continuous, satisfy local Lipschitz conditions in any compact set $X \subset \mathbb{R}^n$, and $f(t,x) \to g(x)$ as $t \to +\infty$, so that the second system is the limit system for the first system. Let $\Phi(t,t_0,x_0)$ and $\varphi(t,t_0,y_0)$ be solutions of these systems, respectively. Suppose that $e \in X$ is a locally asymptotically stable equilibrium of the limit system and its attractive region is

$$W(e) = \{y \in X \mid \varphi(t,t_0,y) \to e, t \to +\infty\}.$$

Let W_Φ be the omega limit set of $\Phi(t,t_0,x_0)$. If $W_\Phi \cap W(e) \neq \varnothing$, then $\lim_{t\to+\infty} \Phi(t,t_0,x_0) = e$.

3. Dynamics of the Deterministic SIQR Model

3.1. The Existence of Equilibrium Points

For a population dynamics system, studying its equilibrium points is the precondition for predicting the development trend of populations within the system.

Theorem 3.1 *System* (1) *has two equilibrium points,* $E_0 = \left(\frac{A}{\mu}, 0, 0, 0\right)$ *for all parameter values and* $E^* = (S^*, I^*, Q^*, R^*)$ *for* $R_0 > 1$, *here* $S^* \in \left(0, \frac{A}{\mu}\right)$, I^*, Q^* *and* $R^* > 0$.

Proof. Summing up all the equations of model (1), we find the following differential equation: $\frac{\mathrm{d}N}{\mathrm{d}t} = A - \mu N - \mu_1 I - \mu_2 Q$. By comparison theorem, we obtain that the solutions of model (1) exist in the region defined by $\Gamma = \left\{(S,I,Q,R) \in \mathbb{R}^4_+ : S + I + Q + R \le \frac{A}{\mu}, S \ge 0, I \ge 0, Q \ge 0, R \ge 0\right\}$. To get the equilibrium points, we set the right-side of equations to be 0,

$$\begin{cases} A - \mu S - \dfrac{\beta SI}{f(S,I)} = 0, \\ \dfrac{\beta SI}{f(S,I)} - (\delta + \gamma + \mu + \mu_1) I = 0, \\ \delta I - (\rho + \mu + \mu_2) Q = 0, \\ \gamma I + \rho Q - \mu R = 0, \end{cases} \tag{4}$$

which yields

$$I = \frac{A - \mu S}{\delta + \gamma + \mu + \mu_1},\ Q = \frac{\delta}{\rho + \mu + \mu_2} I,\ R = \left(\frac{\gamma}{\mu} + \frac{\rho\delta}{\mu(\rho + \mu + \mu_2)}\right) I,$$

$$\frac{\beta S}{f\left(S, \dfrac{A - \mu S}{\delta + \gamma + \mu + \mu_1}\right)} = \delta + \gamma + \mu + \mu_1.$$

If $I=0$, the model (1) has a disease-free equilibrium $E_0=\left(\frac{A}{\mu},0,0,0\right)$ for all parameter values. And we can get the basic reproduction number $R_0=\frac{\beta A}{(\mu+\alpha_1 A)(\delta+\gamma+\mu+\mu_1)}$ by using next generation method. The value R_0 represents the average number of secondary infections when an infected person enters fully susceptible population. If $I\neq 0$, $I=\frac{A-\mu S}{\delta+\gamma+\mu+\mu_1}>0$ implies $S<\frac{A}{\mu}$. Hence, there is no positive equilibrium point if $S>\frac{A}{\mu}$. Now, we consider the function $g(S)$ defined on the interval $\left[0,\frac{A}{\mu}\right]$, where

$$g(S)=\frac{\beta S}{f\left(S,\frac{A-\mu S}{\delta+\gamma+\mu+\mu_1}\right)}-(\delta+\gamma+\mu+\mu_1)$$
$$\triangleq h(S,I)-(\delta+\gamma+\mu+\mu_1).$$

Obviously, $g(0)=-(\delta+\gamma+\mu+\mu_1)<0$ and $g\left(\frac{A}{\mu}\right)=\frac{\beta A}{\mu+\alpha_1 A}-(\delta+\gamma+\mu+\mu_1)=(\delta+\gamma+\mu+\mu_1)(R_0-1)>0$ when $R_0>1$. Simultaneously, differentiating the function *g*, we gain $g'(S)=\frac{\partial h}{\partial S}-\frac{\mu}{\delta+\gamma+\mu+\mu_1}\frac{\partial h}{\partial I}>0$. Because $g(S)$ is monotonically increasing in the interval $\left[0,\frac{A}{\mu}\right]$, $g(0)<0$ and $g\left(\frac{A}{\mu}\right)>0$, the equation $g(S)=0$ has only one positive root by the zero theorem. That is, there exists a unique endemic equilibrium $E^*=(S^*,I^*,Q^*,R^*)$ with $S^*\in\left(0,\frac{A}{\mu}\right)$.

3.2. The Stability of Equilibrium Points

In the biological sense, we analyze the stability of the disease-free equilibrium point and the endemic equilibrium point.

Theorem 3.2. *The disease-free equilibrium* E_0 *of system* (1) *is globally asymptotically stable if* $R_0<1$ *and unstable if* $R_0>1$.

Proof. Consider the Jacobian matrix of system (1) at E_0

$$J(E_0)=\begin{pmatrix}-\mu & -\frac{\beta A}{\mu+\alpha_1 A} & 0 & 0\\ 0 & \frac{\beta A}{\mu+\alpha_1 A}-(\delta+\gamma+\mu+\mu_1) & 0 & 0\\ 0 & \delta & -(\rho+\mu+\mu_2) & 0\\ 0 & \gamma & \rho & -\mu\end{pmatrix}.$$

The characteristic equation of system (1) at E_0 is

$$(\lambda+\mu)^2\left[\lambda+(\rho+\mu+\mu_2)\right]\left[\lambda-\frac{\beta A}{\mu+\alpha_1 A}+(\delta+\gamma+\mu+\mu_1)\right]=0.$$

Clearly, $\lambda_{1,2}=-\mu<0$, $\lambda_3=-(\rho+\mu+\mu_2)<0$ and the positive and negative of the fourth eigenvalue depends on R_0. That is,

$\lambda_4=\dfrac{\beta A}{\mu+\alpha_1 A}-(\delta+\gamma+\mu+\mu_1)=(\delta+\gamma+\mu+\mu_1)(R_0-1)<0$ when $R_0<1$, $\lambda_4>0$ when $R_0>1$. Hence the disease-free equilibrium E_0 is locally asymptotically stable if $R_0<1$ and unstable if $R_0>1$.

Then we prove the global stability of the system (1) at the equilibrium E_0 when $R_0<1$. Taking the Lyapunov function $W_1(t)=I(t)$ into consideration, we get

$$\begin{aligned}\dot{W}_1&=\left(\frac{\beta S}{f(S,I)}-(\delta+\gamma+\mu+\mu_1)\right)I\\&\le\left(\frac{\beta A}{\mu+\alpha_1 A}-(\delta+\gamma+\mu+\mu_1)\right)I\\&=(\delta+\gamma+\mu+\mu_1)(R_0-1)I\le 0.\end{aligned}$$

Thus if $R_0<1$, $\dot{W}_1\le 0$. And $\dot{W}_1=0$ if and only if $I=0$. In this case, $\dfrac{\mathrm{d}S}{\mathrm{d}t}=A-\mu S$ indicates $S\to\dfrac{A}{\mu}$ as $t\to\infty$. Similarly, $Q\to 0$ and $R\to 0$ as $t\to\infty$. So the largest positive invariant set in $\{(S,I,Q,R)\in\Gamma:\dot{W}_1=0\}$ is the singleton E_0. By Liapunov-Lasalle theorem, $E_0=\left(\dfrac{A}{\mu},0,0,0\right)$ is globally asymptotically stable in Γ.

Theorem 3.3. *If* $R_0>1$, *the endemic equilibrium point* E^* *of the system* (1) *is globally asymptotically stable in the region* $\Omega=\Gamma-\{(S,I,Q,R)\in\Gamma:I=0\}$.

Proof. Consider the Jacobian matrix of system (1) at E^*

$$J(E^*)=\begin{pmatrix}-\mu-\dfrac{\beta I^*+\alpha_2\beta I^{*2}}{f^2(S^*,I^*)} & -\dfrac{\beta S^*+\alpha_1\beta S^{*2}}{f^2(S^*,I^*)} & 0 & 0\\ \dfrac{\beta I^*+\alpha_2\beta I^{*2}}{f^2(S^*,I^*)} & \dfrac{\beta S^*+\alpha_1\beta S^{*2}}{f^2(S^*,I^*)}-(\delta+\gamma+\mu+\mu_1) & 0 & 0\\ 0 & \delta & -(\rho+\mu+\mu_2) & 0\\ 0 & \gamma & \rho & -\mu\end{pmatrix}.$$

Let $C_1=\dfrac{\beta I^*+\alpha_2\beta I^{*2}}{f^2(S^*,I^*)}$ and $C_2=\dfrac{\beta S^*+\alpha_1\beta S^{*2}}{f^2(S^*,I^*)}$, then

$$J(E^*)=\begin{pmatrix}-\mu-C_1 & -C_2 & 0 & 0\\ C_1 & C_2-(\delta+\gamma+\mu+\mu_1) & 0 & 0\\ 0 & \delta & -(\rho+\mu+\mu_2) & 0\\ 0 & \gamma & \rho & -\mu\end{pmatrix}.$$

Therefore the characteristic equation of system (1) at E^* is

$$(\lambda+\mu)[\lambda+(\rho+\mu+\mu_2)]\{(\lambda+\mu+C_1)[\lambda+(\delta+\gamma+\mu+\mu_1)-C_2]+C_1C_2\}=0.$$

Obviously, $\lambda_1=-\mu<0$, $\lambda_2=-(\rho+\mu+\mu_2)<0$ and the other two eigenva-

lues are determined by the following quadratic equation

$$\begin{aligned}&\lambda^2+\left[C_1+\mu+\left(\delta+\gamma+\mu+\mu_1\right)-C_2\right]\lambda\\&+\left(C_1+\mu\right)\left[\left(\delta+\gamma+\mu+\mu_1\right)-C_2\right]+C_1C_2=0\\&\Rightarrow\lambda^2+a_1\lambda+a_2=0,\end{aligned}$$

where

$$\begin{aligned}a_1&=C_1+\mu+\left[\left(\delta+\gamma+\mu+\mu_1\right)-C_2\right],\\a_2&=C_1\left(\delta+\gamma+\mu+\mu_1\right)+\mu\left[\left(\delta+\gamma+\mu+\mu_1\right)-C_2\right].\end{aligned}$$

By utilizing Routh-Hurwitz criteria, we know that the system is stable if $a_1,a_2>0$ and unstable if $a_1,a_2<0$. From the second equation of (4), we obtain $\left(\delta+\gamma+\mu+\mu_1\right)>C_2$, thus all eigenvalues have negative real parts. The endemic equilibrium E^* is locally asymptotically stable.

Now we confirm the global stability at the equilibrium E^* when $R_0>1$. The first two equations of system (1) do not contain Q and R, so we consider the following Lyapunov function in the positive quadrant of the two-dimensional plane SI.

$$W_2\left(t\right)=S-S^*-\int_{S^*}^{S}\frac{l\left(S^*,I^*\right)}{l\left(x,I^*\right)}\mathrm{d}x+I^*\Psi\left(\frac{I}{I^*}\right),$$

where $l\left(S,I\right)=\dfrac{\beta S}{f\left(S,I\right)},\Psi\left(x\right)=x-1-\ln x,x>0$. Clearly, $\Psi:\mathbb{R}^+\to\mathbb{R}$ attains its global minimum at $x=1$ and $\Psi\left(1\right)=0$. Besides, the function $\varpi\left(S\right)=S-S^*-\int_{S^*}^{S}\dfrac{l\left(S^*,I^*\right)}{l\left(x,I^*\right)}\mathrm{d}x$ has the global minimum at $S=S^*$ and $\varpi\left(S^*\right)=0$. Then, $\Psi\left(x\right)\geq0$ for any $x>0$ and $\varpi\left(S\right)\geq0$ for any $S>0$. Consequently, $W_2\left(t\right)\geq0$ with equality holding if and only if $\dfrac{S\left(t\right)}{S^*}=\dfrac{I\left(t\right)}{I^*}=1$ for all $t\geq0$. And $l\left(S^*,I^*\right)I^*=\dfrac{\beta S^*I^*}{f\left(S^*,I^*\right)}=\left(\delta+\gamma+\mu+\mu_1\right)I^*$, So the derivative function of $W_2\left(t\right)$ is given by

$$\begin{aligned}\dot{W}_2&=\left(1-\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}\right)\dot{S}+\left(1-\frac{I^*}{I}\right)\dot{I}\\&=\left(1-\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}\right)\left(A-\mu S-\frac{\beta SI}{f\left(S,I\right)}\right)+\left(1-\frac{I^*}{I}\right)\left(\frac{\beta SI}{f\left(S,I\right)}-\left(\delta+\gamma+\mu+\mu_1\right)I\right)\\&=\left(1-\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}\right)\left(-\mu\left(S-S^*\right)+l\left(S^*,I^*\right)I^*-l\left(S,I\right)I\right)\\&\quad+l\left(S^*,I^*\right)I^*\left(1-\frac{I^*}{I}\right)\left(\frac{l\left(S,I\right)I}{l\left(S^*,I^*\right)I^*}-\frac{I}{I^*}\right)\end{aligned}$$

$$
\begin{aligned}
&= \mu\left(S^*-S\right)\left(1-\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}\right)+l\left(S^*,I^*\right)I^*\left(-1-\frac{I}{I^*}+\frac{l\left(S,I^*\right)}{l\left(S,I\right)}+\frac{l\left(S,I\right)I}{l\left(S,I^*\right)I^*}\right)\\
&\quad -l\left(S^*,I^*\right)I^*\left(\frac{l\left(S,I^*\right)}{l\left(S,I\right)}+\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}+\frac{l\left(S,I\right)}{l\left(S^*,I^*\right)}-3\right)\\
&= \mu\left(S^*-S\right)\left(1-\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}\right)+l\left(S^*,I^*\right)I^*\left(1-\frac{l\left(S,I\right)}{l\left(S,I^*\right)}\right)\left(\frac{l\left(S,I^*\right)}{l\left(S,I\right)}-\frac{I}{I^*}\right)\\
&\quad -l\left(S^*,I^*\right)I^*\left(\Psi\left(\frac{l\left(S,I^*\right)}{l\left(S,I\right)}\right)+\Psi\left(\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}\right)+\Psi\left(\frac{l\left(S,I\right)}{l\left(S^*,I^*\right)}\right)\right)
\end{aligned}
$$

Due to

$$\frac{\partial l\left(S,I\right)}{\partial S}>0 \quad \text{for all } S>0 \text{ and } I\geq 0,$$

$$\frac{\partial l\left(S,I\right)}{\partial I}<0 \quad \text{for all } S\geq 0 \text{ and } I\geq 0,$$

we have $\mu\left(S^*-S\right)\left(1-\frac{l\left(S^*,I^*\right)}{l\left(S,I^*\right)}\right)\leq 0$, $\left(1-\frac{l\left(S,I\right)}{l\left(S,I^*\right)}\right)\left(\frac{l\left(S,I^*\right)}{l\left(S,I\right)}-\frac{I}{I^*}\right)\leq 0$, thus $\dot{W}_2\left(t\right)\leq 0$ and $\dot{W}_2=0$ if and only if $S=S^*$ and $I=I^*$. By the Liapunov-Lasalle theorem, all solutions starting in the positive quadrant of SI-plane with $S+I\leq\frac{A}{\mu}$ approach $\left(S^*,I^*\right)$ at $t\to\infty$. In this case, the differential equation for Q has the limiting equation $\frac{\mathrm{d}Q}{\mathrm{d}t}=\delta I^*-\left(\rho+\mu+\mu_2\right)Q$ which implies $Q\to Q^*$ as $t\to\infty$, and similarly, the limiting equation for R is $\frac{\mathrm{d}R}{\mathrm{d}t}=\gamma I^*+\rho Q^*-\mu R$ so that $R\to R^*$ as $t\to\infty$. Therefore, by Lemma 2.3, the endemic equilibrium $E^*=\left(S^*,I^*,Q^*,R^*\right)$ is globally asymptotically stable in the region Ω for the system (1). This completes the proof.

4. Dynamics of the Stochastic SIQR Model

4.1. Existence and Uniqueness of the Global Positive Solution

In order to study the dynamics of stochastic models, the primary question to be considered is whether the solution is global and nonnegative existence. Although the coefficients of the model (2) satisfy the local Lipschitz condition, it's not enough to prove that the solution does not explode within a finite time for any given initial value. Hence in this section, we will show that the solution of model (2) is positive and global.

Theorem 4.1. *For any given initial value* $X\left(0\right)=\left(S\left(0\right),I\left(0\right),Q\left(0\right),R\left(0\right)\right)\in\mathbb{R}_+^4$, *there exists a unique solution* $X\left(t\right)=\left(S\left(t\right),I\left(t\right),Q\left(t\right),R\left(t\right)\right)$ *of system* (2) *on* $t\geq 0$, *which is in* $\mathbb{R}_+^4$ *with probability one.*

Proof. Since the system (2) has locally Lipschitz continuous coefficients, then for any initial value $X(0)\in\mathbb{R}_+^4$, system (2) has a unique local solution $X(t)$ on $[0,\tau_e)$, where τ_e is the explosion time. To verify this solution is global, we only need to show that $\tau_e=\infty$ *a.s.*. Now define the stopping time τ^+ as

$$\tau^+=\inf\left\{t\in[0,\tau_e):\min\left\{S(t),I(t),Q(t),R(t)\right\}\le 0\right\}.$$

Set $\inf\varnothing=\infty$ ($\varnothing$ denotes the empty set). Obviously $\tau^+<\tau_e$, if we can show $\tau^+=\infty$ *a.s.*, then $\tau_e=\infty$ *a.s.*. Assume that this statement is false, then there exists a constant $T>0$ such that $\mathbb{P}\left(\tau^+<T\right)>0$.

Define a C^2-function $V:\mathbb{R}_+^4\to\mathbb{R}$ by

$$V(S,I,Q,R)=\ln(SIQR).$$

Applying Itô's formula, for all $t\in\left[0,\tau^+\right)$, we obtain

$$\mathrm{d}V(S,I,Q,R)\triangleq LV\mathrm{d}t-\sigma_1\mathrm{d}B_1-\sigma_2\mathrm{d}B_2-\sigma_3\frac{I-S}{f(S,I)}\mathrm{d}B_3,$$

where

$$\begin{aligned}LV=&\frac{A}{S}-\mu-\frac{\beta I}{f(S,I)}-\frac{1}{2}\left(\frac{\sigma_3 I}{f(S,I)}\right)^2+\frac{\beta S}{f(S,I)}-(\delta+\gamma+\mu+\mu_1)\\&-\frac{1}{2}\sigma_1^2-\frac{1}{2}\left(\frac{\sigma_3 S}{f(S,I)}\right)^2+\delta\frac{I}{Q}-(\rho+\mu+\mu_2)-\frac{1}{2}\sigma_2^2+\gamma\frac{I}{R}+\rho\frac{Q}{R}-\mu.\end{aligned}$$

Since $S,I,Q,R>0$ for all $t\in\left[0,\tau^+\right)$, and $f(S,I)\ge 1$, we have

$$LV\ge-(4\mu+\delta+\gamma+\mu_1+\rho+\mu_2)-\beta I-\frac{1}{2}\left[\sigma_1^2+\sigma_2^2+(\sigma_3 I)^2+(\sigma_3 S)^2\right].$$

Hence,

$$\begin{aligned}&V(S,I,Q,R)\\&\ge V(S(0),I(0),Q(0),R(0))-\int_0^t\Big[(4\mu+\delta+\gamma+\mu_1+\rho+\mu_2)+\beta I(\theta)\\&\quad+\frac{1}{2}\left(\sigma_1^2+\sigma_2^2+\left[\sigma_3 I(\theta)\right]^2+\left[\sigma_3 S(\theta)\right]^2\right)\Big]\mathrm{d}\theta-\sigma_1 B_1(t)-\sigma_2 B_2(t)\\&\quad-\sigma_3\int_0^t\frac{I(\theta)-S(\theta)}{f(S(\theta),I(\theta))}\mathrm{d}B_3(\theta).\end{aligned}\tag{5}$$

From the definition of τ^+, it follows that $S\left(\tau^+\right)I\left(\tau^+\right)Q\left(\tau^+\right)R\left(\tau^+\right)=0$. Therefore,

$$\lim_{t\to\tau^+}V(S,I,Q,R)=\lim_{t\to\tau^+}\ln(SIQR)=-\infty.$$

Letting $t\to\tau^+$ in (5) and then taking the expectation on both sides of (5), we have that

$$\begin{aligned}-\infty\ge&V(S(0),I(0),Q(0),R(0))-E\int_0^{\tau^+}\Big[(4\mu+\delta+\gamma+\mu_1+\rho+\mu_2)\\&+\beta I(\theta)+\frac{1}{2}\left(\sigma_1^2+\sigma_2^2+\left[\sigma_3 I(\theta)\right]^2+\left[\sigma_3 S(\theta)\right]^2\right)\Big]\mathrm{d}\theta>-\infty,\end{aligned}$$

which is a contradiction and we confirmed $\tau^+ = \infty$ *a.s.*. This completes the proof.

Remark 4.2. *The region* Γ *is almost surely positive invariant of stochastic model* (2), *refer to* [12]. *In addition, from biological consideration, we next focus on the disease dynamics of model* (2) *in the bounded set* Γ.

4.2. The Extinction and Persistent in the Mean of the Disease

One of the most concerning issues in epidemiology is how to establish the threshold condition for the extinction and persistence of the disease. The target of this section is to study the extinction and persistence of the disease. First of all, we define corresponding random threshold as follows:

$$R_0^s = \frac{2\beta A\left(\mu+\alpha_1 A\right)-\sigma_3^2 A^2}{2\left(\mu+\alpha_1 A\right)^2\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)}.$$

Theorem 4.3. *Let* $X(t)=\left(S(t),I(t),Q(t),R(t)\right)$ *be a solution of system* (2) *for any given initial value* $X(0)\in\Gamma$.

1) If $R_0^s<1$ and $\sigma_3^2 \le \frac{\beta\left(\mu+\alpha_1 A\right)}{A}$, then

$$\limsup_{t\to\infty}\frac{\ln I(t)}{t}\le\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\left(R_0^s-1\right)<0 \quad a.s., \tag{6}$$

2) If $\sigma_3^2 > \frac{\beta^2}{2\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)}$, then

$$\limsup_{t\to\infty}\frac{\ln I(t)}{t}\le\frac{\beta^2}{2\sigma_3^2}-\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)<0 \quad a.s., \tag{7}$$

which means that $I(t)$ tends to zero exponentially a.s., *i.e.* the disease dies out with probability 1. Furthermore,

$$\lim_{t\to\infty}\left\langle S(t)\right\rangle=\frac{A}{\mu} \quad a.s.. \tag{8}$$

Proof. Define Lyapunov function $\ln I$, by Itô's formula, we get that

$$\begin{aligned}\mathrm{d}\ln I(t)&=\left[\frac{\beta S}{f(S,I)}-\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)-\frac{1}{2}\sigma_3^2\frac{S^2}{f^2(S,I)}\right]\mathrm{d}t\\&\quad-\sigma_1\mathrm{d}B_1+\sigma_3\frac{S}{f(S,I)}\mathrm{d}B_3\\&=\psi\left(\frac{S}{f(S,I)}\right)\mathrm{d}t-\sigma_1\mathrm{d}B_1+\sigma_3\frac{S}{f(S,I)}\mathrm{d}B_3,\end{aligned} \tag{9}$$

where $\psi(x)=-\frac{1}{2}\sigma_3^2x^2+\beta x-\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)$.

Suppose 1) holds. Noting that $\psi(x)$ is monotone increasing for $x\in\left[0,\frac{\beta}{\sigma_3^2}\right]$ and $\sigma_3^2\le\frac{\beta\left(\mu+\alpha_1 A\right)}{A}$, we have

$$0 \le \frac{S}{f(S,I)} \le \frac{A}{\mu+\alpha_1 A} \le \frac{\beta}{\sigma_3^2}.$$

Then,

$$\mathrm{d}\ln I(t) \le \psi\left(\frac{A}{\mu+\alpha_1 A}\right)\mathrm{d}t - \sigma_1 \mathrm{d}B_1 + \sigma_3 \frac{S}{f(S,I)}\mathrm{d}B_3.$$

Integrating both sides of the above inequality from 0 to t and dividing by t, we obtain

$$\frac{\ln I(t)}{t} \le \frac{1}{t}\int_0^t \psi\left(\frac{A}{\mu+\alpha_1 A}\right)\mathrm{d}\theta + \frac{\ln I(0)}{t} - \sigma_1 \frac{B_1(t)}{t} + \frac{M(t)}{t}, \tag{10}$$

where $M(t) = \sigma_3 \int_0^t \frac{S(\theta)}{f(S(\theta),I(\theta))}\mathrm{d}B_3(\theta)$. By the strong law of large numbers for local martingales [17], we derive that $\lim_{t\to\infty}\frac{M(t)}{t} = \lim_{t\to\infty}\frac{B_1(t)}{t} = 0$ *a.s.*. Since $R_0^S < 1$, Equation (10) becomes

$$\begin{aligned}
&\limsup_{t\to\infty}\frac{\ln I(t)}{t} \le \psi\left(\frac{A}{\mu+\alpha_1 A}\right)\\
&= -\frac{1}{2}\sigma_3^2\left(\frac{A}{\mu+\alpha_1 A}\right)^2 + \beta\left(\frac{A}{\mu+\alpha_1 A}\right) - \left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\\
&= \left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\left[\frac{2\beta A(\mu+\alpha_1 A)-\sigma_3^2 A^2}{2(\mu+\alpha_1 A)^2\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)} - 1\right]\\
&= \left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\left(R_0^s - 1\right) < 0 \quad a.s..
\end{aligned}$$

We obtain the desired assertion (6).

If 2) holds, from Equation (9), we get

$$\begin{aligned}
\frac{\ln I(t)}{t} &= \frac{1}{t}\int_0^t \left[-\frac{1}{2}\sigma_3^2\left(\frac{S(\theta)}{f(S(\theta),I(\theta))} - \frac{\beta}{\sigma_3^2}\right)^2 + \frac{\beta^2}{2\sigma_3^2}\right.\\
&\quad \left. - \left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\right]\mathrm{d}\theta + \frac{\ln I(0)}{t} - \sigma_1\frac{B_1(t)}{t} + \frac{M(t)}{t}\\
&\le \frac{1}{t}\int_0^t \left[\frac{\beta^2}{2\sigma_3^2} - \left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\right]\mathrm{d}\theta\\
&\quad + \frac{\ln I(0)}{t} - \sigma_1\frac{B_1(t)}{t} + \frac{M(t)}{t}.
\end{aligned} \tag{11}$$

Since $\sigma_3^2 > \frac{\beta^2}{2\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)}$, Equation (11) becomes

$$\limsup_{t\to\infty}\frac{\ln I(t)}{t} \le \frac{\beta^2}{2\sigma_3^2} - \left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right) < 0 \quad a.s..$$

We obtain the desired assertion (7). And so

$$\lim_{t\to\infty} I(t) = 0 \quad a.s.. \tag{12}$$

From the first two equations of system (2), there is

$$\mathrm{d}\big(S(t)+I(t)\big) = \big[A-\mu S(t)-(\delta+\gamma+\mu+\mu_1)I(t)\big]\mathrm{d}t - \sigma_1 I \mathrm{d}B_1(t). \tag{13}$$

Integrating both sides of (13) from 0 to t and dividing by t, we have

$$\frac{S(t)+I(t)}{t} - \frac{S(0)+I(0)}{t}$$
$$= A - \frac{\mu}{t}\int_0^t S(\theta)\mathrm{d}\theta - \frac{\delta+\gamma+\mu+\mu_1}{t}\int_0^t I(\theta)\mathrm{d}\theta - \frac{\sigma_1}{t}\int_0^t I(\theta)\mathrm{d}B_1(\theta).$$

Therefore,

$$\langle S(t)\rangle = \frac{A}{\mu} - \frac{\delta+\gamma+\mu+\mu_1}{\mu}\langle I(t)\rangle - \frac{H_1(t)}{t}, \tag{14}$$

where $H_1(t) = \dfrac{S(t)+I(t)}{\mu} - \dfrac{S(0)+I(0)}{\mu} + \dfrac{\sigma_1}{\mu}\int_0^t I(\theta)\mathrm{d}B_1(\theta)$. Clearly,

$\lim_{t\to\infty} \dfrac{H_1(t)}{t} = 0$ *a.s.* and from (12), we have

$$\lim_{t\to\infty}\langle S(t)\rangle = \frac{A}{\mu} - \frac{\delta+\gamma+\mu+\mu_1}{\mu}\lim_{t\to\infty}\langle I(t)\rangle - \lim_{t\to\infty}\frac{H_1(t)}{t} = \frac{A}{\mu} \quad a.s..$$

Therefore the assertion (8) holds. The conclusion is proven.

Next, the conditions for the persistence of the disease are presented.

Theorem 4.4. *Suppose that* $R_0^S > 1$, *then the solution* $X(t) = \big(S(t), I(t), Q(t), R(t)\big)$ *of system* (2) *is persistent in the mean for any given initial value* $X(0) \in \Gamma$. *Moreover*,

$$\liminf_{t\to\infty}\langle I(t)\rangle \geq I_* > 0, \tag{15}$$

$$\liminf_{t\to\infty}\left\langle \frac{A}{\mu} - S(t)\right\rangle \geq \frac{\delta+\gamma+\mu+\mu_1}{\mu} I_* > 0, \tag{16}$$

$$\liminf_{t\to\infty}\langle Q(t)\rangle \geq \frac{\delta}{\rho+\mu+\mu_2} I_* > 0, \tag{17}$$

$$\liminf_{t\to\infty}\langle R(t)\rangle \geq \frac{\gamma(\rho+\mu+\mu_2)+\rho\delta}{\mu(\rho+\mu+\mu_2)} I_* > 0, \tag{18}$$

where

$$I_* = \frac{\mu(\mu+\alpha_1 A)\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\left(R_0^S-1\right)}{\beta\big[\mu(\delta+\gamma+\mu+\mu_1)+A(\alpha_2\mu+\alpha_3 A)\big]}.$$

Proof. Since $X \in \Gamma$, we have

$$\frac{\beta S}{f(S,I)} = \frac{\beta A}{\mu+\alpha_1 A} - \frac{\beta\mu}{f(S,I)(\mu+\alpha_1 A)}\left(\frac{A}{\mu}-S\right)$$
$$- \frac{\beta\alpha_2 A}{f(S,I)(\mu+\alpha_1 A)} I - \frac{\beta\alpha_3 A}{f(S,I)(\mu+\alpha_1 A)} SI$$

$$\geq \frac{\beta A}{\mu+\alpha_1 A}-\frac{\beta\mu}{\mu+\alpha_1 A}\left(\frac{A}{\mu}-S\right)-\frac{\beta\alpha_2 A}{\mu+\alpha_1 A}I-\frac{\beta\alpha_3 A}{\mu+\alpha_1 A}SI$$
$$\geq \frac{\beta\mu S}{\mu+\alpha_1 A}-\frac{\beta A}{\mu+\alpha_1 A}\left(\alpha_2+\alpha_3\frac{A}{\mu}\right)I.$$

Then

$$\begin{aligned}\mathrm{d}\ln I(t) &\geq \left[\frac{\beta S}{f(S,I)}-\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)-\frac{1}{2}\sigma_3^2\frac{A^2}{(\mu+\alpha_1 A)^2}\right]\mathrm{d}t\\ &\quad -\sigma_1\mathrm{d}B_1+\sigma_3\frac{S}{f(S,I)}\mathrm{d}B_3\\ &\geq \left[\frac{\beta\mu S}{\mu+\alpha_1 A}-\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)-\frac{\sigma_3^2A^2}{2(\mu+\alpha_1 A)^2}\right]\mathrm{d}t\\ &\quad -\frac{\beta A}{\mu+\alpha_1 A}\left(\alpha_2+\alpha_3\frac{A}{\mu}\right)I\mathrm{d}t-\sigma_1\mathrm{d}B_1+\sigma_3\frac{S}{f(S,I)}\mathrm{d}B_3.\end{aligned}\tag{19}$$

Integrating both sides of (19) from 0 to t, there is

$$\begin{aligned}&\ln I(t)-\ln I(0)\\ &\geq \frac{\beta\mu}{\mu+\alpha_1 A}\int_0^t S(\theta)\mathrm{d}\theta-\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)t-\frac{\sigma_3^2A^2}{2(\mu+\alpha_1 A)^2}t\\ &\quad -\frac{\beta A}{\mu+\alpha_1 A}\left(\alpha_2+\alpha_3\frac{A}{\mu}\right)\int_0^t I(\theta)\mathrm{d}\theta-\sigma_1B_1(t)+M(t).\end{aligned}$$

From (14), we have

$$\begin{aligned}\ln I(t) &\geq \left[\frac{\beta A}{\mu+\alpha_1 A}-\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)-\frac{\sigma_3^2A^2}{2(\mu+\alpha_1 A)^2}\right]t\\ &\quad -\frac{\beta}{\mu+\alpha_1 A}\left[(\delta+\gamma+\mu+\mu_1)+A\left(\alpha_2+\alpha_3\frac{A}{\mu}\right)\right]\int_0^t I(\theta)\mathrm{d}\theta+H_2(t)\\ &=\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\left(R_0^S-1\right)t\\ &\quad -\frac{\beta}{\mu+\alpha_1 A}\left[(\delta+\gamma+\mu+\mu_1)+A\left(\alpha_2+\alpha_3\frac{A}{\mu}\right)\right]\int_0^t I(\theta)\mathrm{d}\theta+H_2(t),\end{aligned}$$

where $H_2(t)=\ln I(0)-\sigma_1B_1(t)+M(t)-\frac{\beta\mu}{\mu+\alpha_1 A}H_1(t)$. Obviously, $\lim_{t\to\infty}\frac{H_2(t)}{t}=0$ *a.s.*. By Lemma 2.2 and $R_0^S>1$, we deduce that

$$\liminf_{t\to\infty}\langle I(t)\rangle \geq \frac{\mu(\mu+\alpha_1 A)\left(\delta+\gamma+\mu+\mu_1+\frac{1}{2}\sigma_1^2\right)\left(R_0^S-1\right)}{\beta\left[\mu(\delta+\gamma+\mu+\mu_1)+A(\alpha_2\mu+\alpha_3 A)\right]}=I_*>0.$$

This is the required inequality (15), and from (14), we have

$$\left\langle\frac{A}{\mu}-S(t)\right\rangle=\frac{\delta+\gamma+\mu+\mu_1}{\mu}\langle I(t)\rangle+\frac{H_1(t)}{t}.$$

Therefore,

$$\liminf_{t\to\infty}\left\langle \frac{A}{\mu}-S(t)\right\rangle = \frac{\delta+\gamma+\mu+\mu_1}{\mu}\liminf_{t\to\infty}\langle I(t)\rangle \geq \frac{\delta+\gamma+\mu+\mu_1}{\mu}I_* > 0,$$

the inequality (16) is valid. From the third equation of system (2), we have

$$Q(t)-Q(0)=\delta\int_0^t I(\theta)\mathrm{d}\theta-(\rho+\mu+\mu_2)\int_0^t Q(\theta)\mathrm{d}\theta-\sigma_2\int_0^t Q(\theta)\mathrm{d}B_2(\theta).$$

Then,

$$\langle Q(t)\rangle = \frac{\delta}{\rho+\mu+\mu_2}\langle I(t)\rangle + \frac{H_3(t)}{t},$$

where $H_3(t)=\dfrac{Q(0)-Q(t)}{\rho+\mu+\mu_2}-\dfrac{\sigma_2}{\rho+\mu+\mu_2}\int_0^t Q(\theta)\mathrm{d}B_2(\theta)$. It follows from the strong law of large numbers for local martingales that $\lim_{t\to\infty}\dfrac{H_3(t)}{t}=0$ *a.s.*, hence (17) holds for

$$\liminf_{t\to\infty}\langle Q(t)\rangle = \frac{\delta}{\rho+\mu+\mu_2}\liminf_{t\to\infty}\langle I(t)\rangle \geq \frac{\delta}{\rho+\mu+\mu_2}I_* > 0.$$

The last equation of system (2) gives

$$R(t)-R(0)=\gamma\int_0^t I(\theta)\mathrm{d}\theta+\rho\int_0^t Q(\theta)\mathrm{d}\theta-\mu\int_0^t R(\theta)\mathrm{d}\theta.$$

Then,

$$\langle R(t)\rangle = \frac{\gamma}{\mu}\langle I(t)\rangle + \frac{\rho}{\mu}\langle Q(t)\rangle + \frac{R(0)-R(t)}{\mu t}.$$

So we have

$$\liminf_{t\to\infty}\langle R(t)\rangle = \frac{\gamma}{\mu}\liminf_{t\to\infty}\langle I(t)\rangle + \frac{\rho}{\mu}\liminf_{t\to\infty}\langle Q(t)\rangle \geq \frac{\gamma}{\mu}I_* + \frac{\rho\delta}{\mu(\rho+\mu+\mu_2)}I_* = \frac{\gamma(\rho+\mu+\mu_2)+\rho\delta}{\mu(\rho+\mu+\mu_2)}I_* > 0.$$

This is the required inequality (18).

4.3. Numerical Simulations

In this section, we numerically simulate solutions of the models by using the Milstein's method [18] to confirm main results. We compare the threshold parameters of the deterministic model and stochastic model to illustrate the effect of white noise on the system. The model (2) can be rewritten as the following discrete equation:

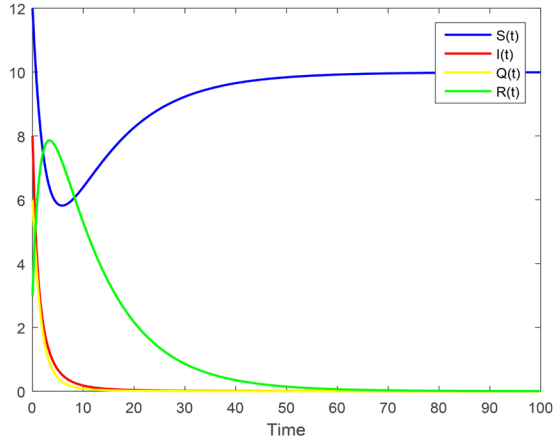
$$
\begin{cases}
S_{k+1} = S_k + \left(A - \mu S_k - \dfrac{\beta S_k I_k}{1+\alpha_1 S_k + \alpha_2 I_k + \alpha_3 S_k I_k} \right) \Delta t \\
\qquad - \dfrac{\sigma_3 S_k I_k}{1+\alpha_1 S_k + \alpha_2 I_k + \alpha_3 S_k I_k} \sqrt{\Delta t} \varepsilon_k \\
\qquad - \dfrac{\sigma_3^2}{2} \dfrac{S_k I_k}{1+\alpha_1 S_k + \alpha_2 I_k + \alpha_3 S_k I_k} \left(\varepsilon_k^2 - 1 \right) \Delta t, \\
I_{k+1} = I_k + \left(\dfrac{\beta S_k I_k}{1+\alpha_1 S_k + \alpha_2 I_k + \alpha_3 S_k I_k} - \left(\delta + \gamma + \mu + \mu_1 \right) I_k \right) \Delta t \\
\qquad - \sigma_1 I_k \sqrt{\Delta t} \eta_k - \dfrac{\sigma_1^2}{2} I_k \left(\eta_k^2 - 1 \right) \Delta t + \dfrac{\sigma_3 S_k I_k}{1+\alpha_1 S_k + \alpha_2 I_k + \alpha_3 S_k I_k} \sqrt{\Delta t} \varepsilon_k \\
\qquad + \dfrac{\sigma_3^2}{2} \dfrac{S_k I_k}{1+\alpha_1 S_k + \alpha_2 I_k + \alpha_3 S_k I_k} \left(\varepsilon_k^2 - 1 \right) \Delta t, \\
Q_{k+1} = Q_k + \left(\delta I_k - \left(\rho + \mu + \mu_2 \right) Q_k \right) \Delta t - \sigma_2 Q_k \sqrt{\Delta t} \xi_k - \dfrac{\sigma_2^2}{2} Q_k \left(\xi_k^2 - 1 \right) \Delta t, \\
R_{k+1} = R_k + \left(\gamma I_k + \rho Q_k - \mu R_k \right) \Delta t,
\end{cases}
$$

where ε_k, η_k and ξ_k, $k = 1, 2, \cdots, n$ are the Gaussian random variables $N(0,1)$. Similarly, the model (1) can also be written in the above form. We just need to delete the disturbance term and will not repeat it here.

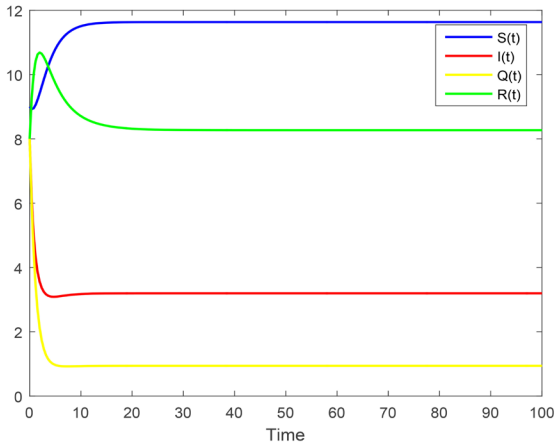
Example 4.3.1 For the deterministic system (1), we choose the initial value $(S(0), I(0), Q(0), R(0)) = (12, 8, 6, 3)$ and the parameter values $A = 1$, $\beta = 0.4$, $\delta = 0.25$, $\gamma = 0.4$, $\rho = 0.4$, $\mu = 0.1$, $\mu_1 = 0.3$, $\mu_2 = 0.25$, $\alpha_1 = 0.3$, $\alpha_2 = 0.1$, $\alpha_3 = 0.08$. By Matlab software, we get $R_0 = 0.9524 < 1$ and find that the class I, Q and R tend to 0, which means that the disease dies out (see **Figure 1(a)**). Theorem 3.2 is illustrated. Then, let $(S(0), I(0), Q(0), R(0)) = (9, 8, 8, 8)$ and $A = 6$, $\beta = 0.4$, $\delta = 0.25$, $\gamma = 0.4$, $\rho = 0.4$, $\mu = 0.2$, $\mu_1 = 0.3$, $\mu_2 = 0.25$, $\alpha_1 = 0.03$, $\alpha_2 = 0.03$, $\alpha_3 = 0.07$ to draw **Figure 1(b)**. It shows that the disease becomes endemic and $R_0 = 5.4920 > 1$. The condition of theorem 3.3 is satisfied.

Example 4.3.2 For the stochastic system (2), we choose the initial value $(S(0), I(0), Q(0), R(0)) = (5, 2, 1, 1)$ and the parameter values $A = 1$, $\beta = 0.6$, $\delta = 0.3$, $\gamma = 0.3$, $\rho = 0.3$, $\mu = 0.1$, $\mu_1 = 0.2$, $\mu_2 = 0.15$, $\alpha_1 = 0.5$, $\alpha_2 = 0.1$, $\alpha_3 = 0.2$, $\sigma_1 = 0.01$, $\sigma_2 = 0.01$, $\sigma_3 = 0.3$. By calculation, $R_0^S = 0.9722 < 1$ and $\sigma_3^2 = 0.09 < 0.36 = \dfrac{\beta(\mu + \alpha_1 A)}{A}$. Hence, the condition 1) of theorem 4.3 is satisfied. In **Figure 2(a)**, the class I exponentially decays to zero which indicates the extinction of the disease. Next, we let parameter $\sigma_3 = 0.7$ and others are the same as above. In this case, $\sigma_3^2 = 0.49 > 0.36 = \dfrac{\beta^2}{2\left(\delta + \gamma + \mu + \mu_1 + \frac{1}{2}\sigma_1^2 \right)}$. Therefore, the condition 2) of theorem 4.3 is satisfied and the disease dies out (**Figure 2(b)**). Finally, let $(S(0), I(0), Q(0), R(0)) = (5, 4, 4, 4)$, $A = 2$, $\alpha_1 = 0.2$, $\sigma_2 = 0.1$, $\sigma_3 = 0.09$

and keep the other parameters, we get $R_0^S = 2.5945 > 1$. According to theorem 4.4, all classes of the system (2) are persistent and are shown in Figure 2(c).

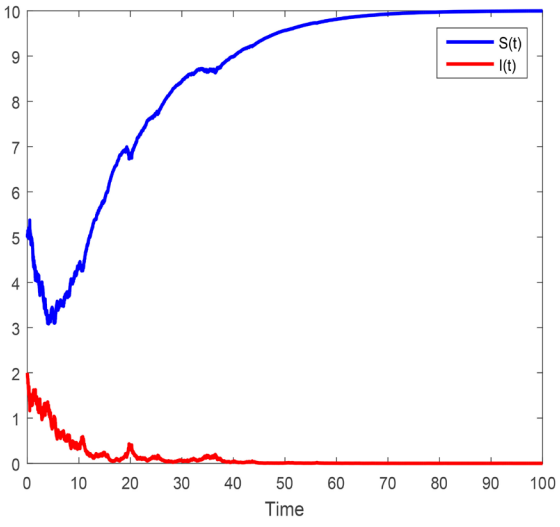


(a)

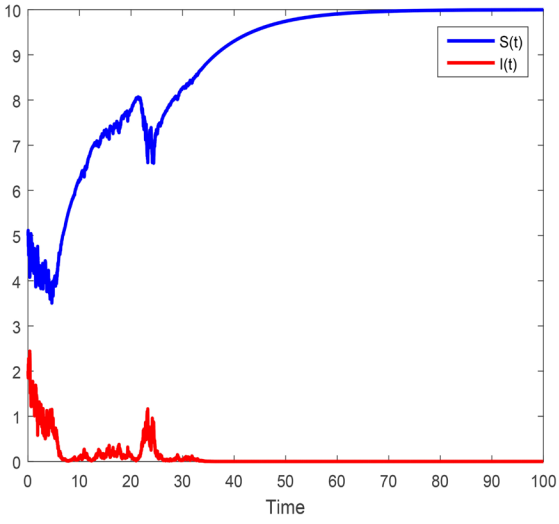


(b)

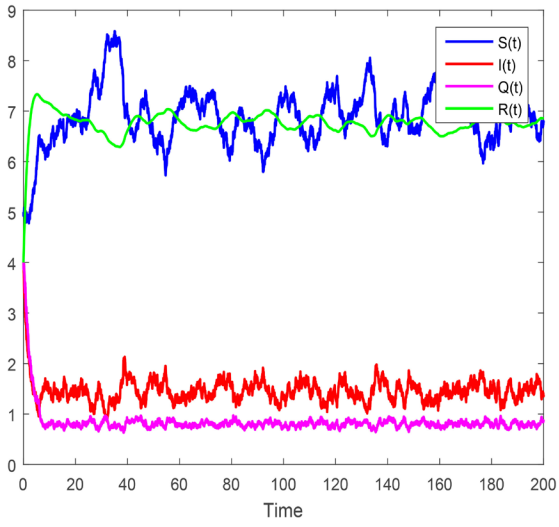
Figure 1. Dynamics of the deterministic system (1).



(a)



(b)



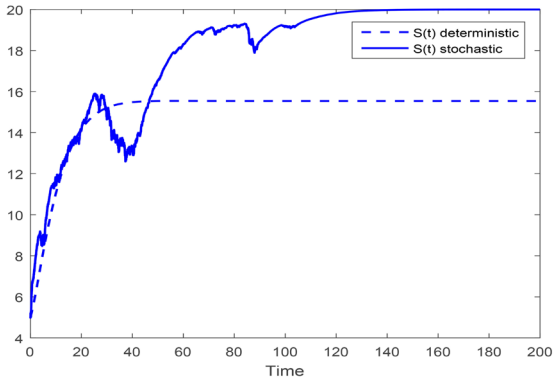
(c)

Figure 2. Dynamics of the stochastic system (2).

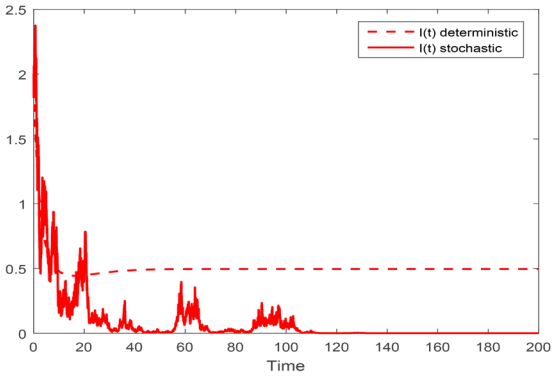
Example 4.3.3 Now, we reselect the parameters $(S(0),I(0),Q(0),R(0))=(5,2,2,2)$, $A=2$, $\beta=0.6$, $\delta=0.3$, $\gamma=0.3$, $\rho=0.3$, $\mu=0.1$, $\mu_1=0.2$, $\mu_2=0.15$, $\alpha_1=0.5$, $\alpha_2=0.1$, $\alpha_3=0.2$, $\sigma_1=0.01$, $\sigma_2=0.01$, $\sigma_3=0.4$ and give a set of comparison charts of simulation results. In **Figure 3**, $R_0^S=0.9182<1<1.2121=R_0$, the class S, I, Q and R of deterministic model all exist, which means that the disease break out, but after adding white noise, except for the class S, the others tend to be 0. It reveals that the random fluctuations can suppress disease prevail.

5. Summary and Discussions

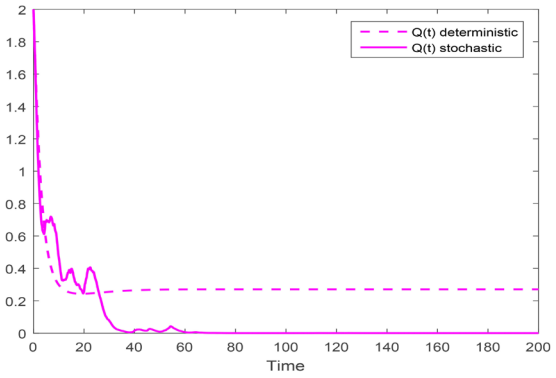
In this work, we investigate the deterministic and stochastic SIQR epidemic models with the specific nonlinear incidence. This incidence rate can become multiple types, and is more abundant than saturation incidence. We obtain the



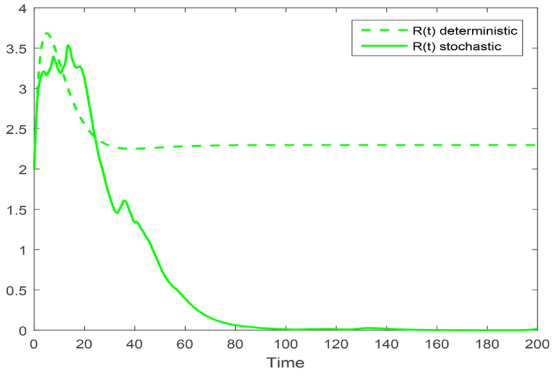
(a)



(b)



(c)



(d)

Figure 3. The random fluctuations can suppress disease break out.

dynamics properties of the SIQR model based on two threshold parameters R_0 and R_0^S. And owing to $R_0^S < R_0$, there may be an interesting situation $R_0^S < 1 < R_0$, which indicates that the random fluctuations can suppress disease break out. Moreover, we simulate them with computer software and the results of the simulation are also consistent with the theoretical results. It can provide us with some useful control strategies to regulate disease dynamics.

In future work, we will further consider the delayed SIQR model with this incidence and the SIQS model without permanent immunity.

Conflicts of Interest

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

References

[1] Jiang, D.Q., Yu, J.J. and Ji, C.Y. (2011) Asymptotic Behavior of Global Positive Solution to a Stochastic SIR Model. *Mathematical and Computer Modelling*, **54**, 221-232. https://doi.org/10.1016/j.mcm.2011.02.004

[2] Hattaf, K., Lashari, A.A., Louartassi, Y. and Yousfi, N. (2013) A Delayed SIR Epidemic Model with General Incidence Rate. *Electronic Journal of Qualitative Theory of Differential Equations*, **3**, 1-9. https://doi.org/10.14232/ejqtde.2013.1.3

[3] Fan, K.G., Zhang, Y., Gao, S.J. and Wei, X. (2017) A Class of Stochastic Delayed SIR Epidemic Models with Generalized Nonlinear Incidence Rate and Temporary Immunity. *Physica A*: *Statistical Mechanics and Its Applications*, **481**, 198-208. https://doi.org/10.1016/j.physa.2017.04.055

[4] Alakes, M., Prosenjit, S. and Samanta, G.P. (2018) Analysis of an SIQR Model. *Journal of Ultra Scientist of Physical Sciences*, **30**, 218-226. https://doi.org/10.22147/jusps-A/300307

[5] Lan, G.J., Chen, Z.W., Wei, C.J. and Zhang, S.W. (2018) Stationary Distribution of a Stochastic SIQR Epidemic Model with Saturated Incidence and Degenerate Diffusion. *Physica A*: *Statistical Mechanics and Its Applications*, **511**, 61-77. https://doi.org/10.1016/j.physa.2018.07.041

[6] Chahrazed, L. and Lazhar, R.F. (2013) Stability of a Delayed SIQRS Model with Temporary Immunity. *Advances in Pure Mathematics*, **3**, 240-245. https://doi.org/10.4236/apm.2013.32034

[7] Zhang, X.B. and Huo, H.F. (2014) Dynamics of the Deterministic and Stochastic SIQS Epidemic Model with Nonlinear Incidence. *Applied Mathematics and Computation*, **243**, 546-558. https://doi.org/10.1016/j.amc.2014.05.136

[8] Joshi, H. and Sharma, R.K. (2017) Global of an SIQR Epidemic Model with Saturated Incidence Rate. *Asian Journal of Mathematics and Computer Research*, **21**, 156-166.

[9] Adnani, J., Hattaf, K. and Yousfi, N. (2013) Stability Analysis of a Stochastic SIR Epidemic Model with Specific Nonlinear Incidence Rate. *International Journal of Stochastic Analysis*, **2013**, Article ID: 431257. https://doi.org/10.1155/2013/431257

[10] Ji, C.Y. and Jiang, D.Q. (2011) Dynamics of a Stochastic Density Dependent Predator-Prey System with Beddington-DeAngelis Functional Response. *Journal of Mathematical Analysis and Applications*, **381**, 441-453. https://doi.org/10.1016/j.jmaa.2011.02.037

[11] Crowley, P.H. and Martin, E.K. (1989) Functional Responses and Interference within and between Year Classes of a Dragonfly Population. *Freshwater Science*, **8**, 211-221. https://doi.org/10.2307/1467324

[12] Adnani, J., Hattaf, K. and Yousfi, N. (2016) Analysis of a Stochastic SIRS Epidemic Model with Specific Functional Response. *Applied Mathematical Sciences*, **10**, 301-314. https://doi.org/10.12988/ams.2016.511697

[13] Hattaf, K., Mahrouf, K. and Adnani, J. (2018) Qualitative Analysis of a Stochastic Epidemic Model with Specific Functional Response and Temporary Immunity. *Physica A*: *Statistical Mechanics and its Applications*, **490**, 591-600. https://doi.org/10.1016/j.physa.2017.08.043

[14] Li, D., Cui, J.A., Liu, M. and Liu, S.Q. (2015) The Evolutionary Dynamics of Stochastic Epidemic Model with Nonlinear Incidence Rate. *Bulletin of Mathematical Biology*, **77**, 1705-1743. https://doi.org/10.1007/s11538-015-0101-9

[15] Cai, Y.L., Kang, Y. and Wang, W.M. (2017) A Stochastic SIRS Epidemic Model with Nonlinear Incidence Rate. *Applied Mathematics and Computation*, **305**, 221-240. https://doi.org/10.1016/j.amc.2017.02.003

[16] Ji, C.Y. and Jiang, D.Q. (2014) Threshold Behaviour of a Stochastic SIR Model. *Applied Mathematical Modelling*, **38**, 5067-5079. https://doi.org/10.1016/j.apm.2014.03.037

[17] Chang, Z.B., Meng, X.Z. and Lu, X. (2017) Analysis of a Novel Stochastic SIRS Epidemic Model with Two Different Saturated Incidence Rates. *Physica A*: *Statistical Mechanics and Its Applications*, **472**, 103-116. https://doi.org/10.1016/j.physa.2017.01.015

[18] Higham, D.J. (2001) An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *Society for Industrial and Applied Mathematics*, **43**, 525-546. https://doi.org/10.1137/S0036144500378302

Journal of Applied Mathematics and Physics, 2019, 7, 1861-1869
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

A Simulation of Transit Timing Variation

Ziyue Zeng

Guangdong Experimental High School, Guangzhou, China
Email: cis2018@foxmail.com

How to cite this paper: Zeng, Z.Y. (2019) A Simulation of Transit Timing Variation. *Journal of Applied Mathematics and Physics*, **7**, 1861-1869.
https://doi.org/10.4236/jamp.2019.78127

Received: June 14, 2019
Accepted: August 20, 2019
Published: August 23, 2019

Abstract

Exoplanet transit timing variation is a method to find exoplanets. To understand this method better, I wrote a computer program in python to simulate the transit of exoplanets. I use my program to simulate the transit timing variation observed in the Kepler-19 system. I make a simple simulation of Kepler-19 system, and this simulation shows that the variation in transit timing due to other planets is very obvious for Kepler-19b, the transiting planet, which means the transit timing variation method is very useful for finding exoplanet in Kepler-19 system. The whole paper is an illustration for that. The simulation I make is relatively simple but it still corresponds to the law of TTV, and because of its simplicity, it can help more people understand.

Keywords

Exoplanet, Transit Timing Variation, Transit Timing, Kepler-19 System

1. Introduction

This paper presents a program that relies primarily on transit photometry and transit timing in the search for exoplanets to simulate the transit timing variation observed in the Kepler-19 system. Transit timing variation is a phenomenon that the period of the planet transiting the star varies due to the gravitational effects of other planets orbiting the star. And we can use this phenomenon to discover other planets which don't transit base on the transit photometry method, which means that the light curve of a star can indicate the existence of an exoplanet that does not transit. There are also other methods to discover exoplanet indirectly, and transit timing variation and other methods can check each other to make sure an exoplanet really exists. So far, the transit timing variation method has only discovered two exoplanets, Kepler-19c [1] and Kepler-9d. Before 2012, the radial-velocity method was the most productive technique for finding exoplanets. After 2012, the transit method becomes the most productive

because of the use of the Kepler spacecraft. But due to their limitations, many exoplanets will be ignored. For example, if the orbital plane of an exoplanet is perpendicular to the line connecting the star surrounded by this exoplanet to the Earth, then we will not be able to observe the change in radial velocity and cannot directly know the existence of this exoplanet through transit, which means both radial-velocity method and transit photometry method don't work in this case. However, transit timing variation can solve this problem and that's why I am so interested in it. TTV is very useful as a supplement and I want to know about its feasibility and want to know whether it can do more than being a supplement. What's more, I hope that more people can know more about it. I wrote a computer program to simulate the transit timing variation of Kepler-19b to better understand the system. The program uses the initial velocity to update force and acceleration, and then uses force and acceleration to update positioning. Current positioning then updates the velocity, and the process repeats. I calculated the initial velocity using Kepler's Second Law and Kepler's Third Law and I calculated the difference between periods to make a period variation graph to show the transit timing variation of Kepler-19b due to Kepler-19c and Kepler-19d. I use geometry to calculate the area of the star that is covered by the transiting planet, and this can be used to draw a light curve of Kepler-19.

2. Simulations

2.1. Updating the Position with an Initial Velocity

Following are the equations I use to update the position of the first planet.

$$v_{\text{final}} = v_{\text{initial}} + a \times t \tag{1}$$

$$v_{\text{final}} = v_{\text{initial}} + \frac{F \times t}{m} \tag{2}$$

$$x_{\text{final}} = x_{\text{initial}} + v_{\text{final}} \times t \tag{3}$$

By expressing and calculating velocity, force, and position in terms of x, y, and z coordinates, we can plot them in a three-dimensional coordinate system. In order to use Equation (2) above to update the velocity, we need to calculate the forces between the planet and the star using Newton's Gravitational law. G is the gravitational constant, M is the mass of the star, and m is the mass of the planet.

$$\vec{F} = \frac{G \times M \times m}{r^2} \hat{r} \tag{4}$$

To use Equation (4) and to express and calculate force in terms of x, y, and z direction, we need to calculate the Euclidean distances in x, y, and z direction.

The force in x direction is given by

$$F_x = F \times \frac{r_x}{r} \times \hat{r}$$

Similarly, the force in y and z direction can be calculated using the method above.

As a result, we can calculate the velocity in terms of x, y, and z direction and

then update the position in terms of *x*, *y*, and *z* direction.

We can use the new position to calculate the new forces between celestial bodies and then use the new forces to update the velocity and position in *x*, *y*, and *z* direction again.

In this process, knowing the mass of the celestial bodies, the only variation that I need is the initial velocity and the initial Euclidean distance between the two celestial bodies.

All the methods mentioned above allow us to simulate the motion of many celestial bodies. To get a graph about the motion of many celestial bodies, we can record every position that the celestial bodies pass through.

I randomly set some properties of celestial bodies to check whether my simulation can work (see Table 1).

In order to make the planets orbit around the star, we need to know something about the initial velocity:

If

$$v_{\text{initial}} = \sqrt{\frac{GM}{r}} \times \hat{r},$$

planets will have a circular orbit.

If

$$\sqrt{\frac{GM}{r}} \times \hat{r} < v_{\text{initial}} < \sqrt{\frac{2GM}{r}} \times \hat{r},$$

planets will have an elliptical orbit.

If

$$v_{\text{initial}} = \sqrt{\frac{2GM}{r}} \times \hat{r},$$

planets will have a hyperbolic orbit and it will escape from the gravitational field of the star.

The formulas for explanation are as follows:

The orbital velocity:

$$F_c = F_g$$

$$\frac{mv_1^2}{r} = \frac{GMm}{r^2}$$

$$v_1 = \sqrt{\frac{GM}{r}}$$

The escape velocity:

$$KE = U$$

$$\frac{1}{2}mv_2^2 = \frac{GMm}{r}$$

$$v_2 = \sqrt{\frac{2GM}{r}}$$

And then I make a graph to check.

Table 1. Some random data of celestial bodies. Celestial body 1 is the star. Initial velocities in y and z directions are 0.

Name	Mass (kg)	Orbital radius (m)	initial velocity (m/s) in x direction (see Figure 1)
celestial body 1	5.95e24	0	−12
celestial body 2	1e20	6,371,000 × 2	7900
celestial body 3	3e21	6,371,000	3000
celestial body 4	7e21	7,892,344	5000
celestial body 5	9e22	10,376,476	7000

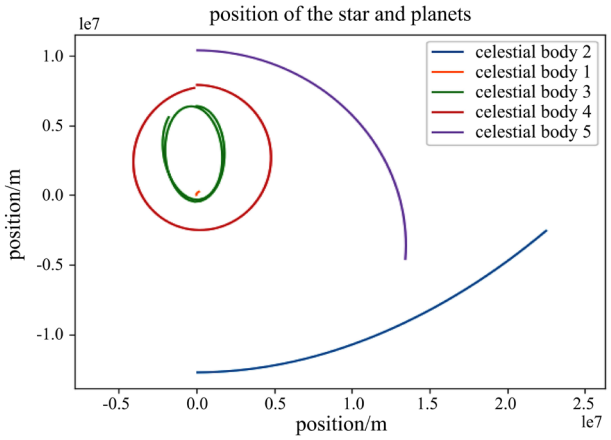


Figure 1. The application of the process describe above.

2.2. Using Maths to Calculate the Initial Velocity

In reality, most of the orbits of celestial bodies are ellipses instead of circle. And the central celestial bodies will at one focus of the ellipse. To calculate the initial velocity of the celestial bodies, we need to use some formulas of ellipse.

The semi-major axis is a, the semi-minor axis is b, and the focal length is c.

Eccentricity:

$$e = \frac{c}{a}$$

semi-minor axis:

$$b^2 = a^2 - c^2$$

Area of ellipse:

$$\text{area} = \pi ab$$

For a line between the star and the planet, the line will sweep out an area as the planet moves around the star for a small angle $\mathrm{d}\theta$ the area will be:

$$\mathrm{d}A = \frac{1}{2} \times (a-c)^2 \times \mathrm{d}\theta$$

As a result, the initial velocity can be calculated by combining the formulas above using Kepler's second law:

$$\frac{\mathrm{d}A}{\mathrm{d}t}=\frac{1}{2}\times\left(a-c\right)^{2}\times\frac{\mathrm{d}\theta}{\mathrm{d}t}=\frac{\pi ab}{\text{period}}$$

$$\frac{1}{2}\times\left(a-c\right)\times v_{\text{initial}}=\frac{\pi ab}{\text{period}}$$

$$v_{\text{initial}}=\frac{2\pi a\times\sqrt{a^{2}\times\left(1-e^{2}\right)}}{\text{period}\times\left(a-e\times a\right)}$$

Actually, we can apply all the things mentioned in this section on celestial bodies which have circular orbit. The eccentricity of a circle is 0 and the semi-major axis is the radius of a circle. The formula of calculating the initial velocity of a circular-orbit planet will be:

$$v_{\text{initial}}=\frac{2\times\pi\times\text{radius}\times\sqrt{\text{radius}^{2}\times\left(1-0^{2}\right)}}{\text{period}\times\left(\text{radius}-0\times a\right)}$$

$$v_{\text{initial}}=\frac{2\times\pi\times\text{raidus}}{\text{period}}=\frac{\text{perimeter}}{\text{time}}$$

So we only need the semi-major axis, eccentricity, and the period to calculate the initial velocity.

2.3. Period Variation

The period of a planet is not constant because of the gravitational effects due to other planets. The period variation is one kind of transit timing variation.

In order to know the differences between periods, I plot periods in a graph in a relatively long time. The way I calculate one period for a particular planet is to record every time when one of the planet passes from negative to positive. And then I calculate the differences between every two record time, which show us the periods of the planets. I randomly choose some data to draw the graph. See **Table 2**.

2.4. Drawing a Light Curve

To draw a light curve of the star, we need the relationship between the area blocked by the transiting planet and the surface area of the star that we can observe when it is not covered by the planet. To calculate the relationship between these two areas, we need to establish a two-dimensional system and put the center of the star at the origin. Let the distance between the center of the planet and the center of the star be d (see **Figure 2**) $d=\text{position}_{\text{star}}-\text{position}_{\text{planet}}$. Because the radius of the orbit of the planet is relatively small compared to the distance between the whole system and the observer, we can let the radius of the planet be R_p and let the radius of the star be R_s. Also because the arc is relatively small compared to the perimeter of the orbit, we can consider the velocity that planet moves toward the star as the linear velocity of the planet. The method of calculating the star's area that is blocked by the planet is changing with respect to the change in the relative position of the planet and the star.

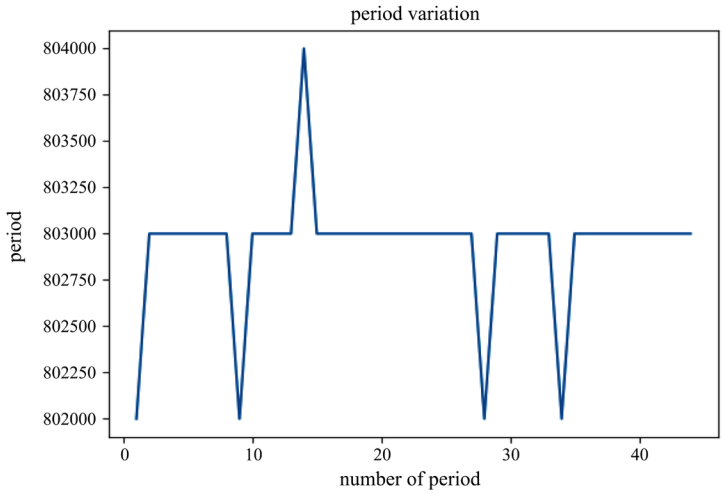


Figure 2. Period variation of a planet due to the gravitational effects of other planets.

Table 2. Some random properties of celestial bodies. Celestial body 1 is the star.

Name	Mass(kg)	eccentricity	the average period
celestial body 1	0.936 × mass of the sun	0	0
celestial body 2	8.4 × mass of the earth	0.12	9.28716
celestial body 3	31 × mass of the earth	0.21	28.731
celestial body 4	55 × mass of the earth	0.05	62.95

There are six conditions that need different methods to calculate the blocked area shown in **Figure 3**.

1) when the planet is not blocking the star: area = 0;

2) when $d \le -\left(R_s + R_p\right)$ and $d \ge -R_s$;

3) when $d < -R_s$ and $d \ge d > -\left(R_s - R_p\right)$;

4) when $d > -\left(R_s - R_p\right)$ and $d \le R_s - R_p$;

5) when $d > R_s - R_p$ and $d \le R_s$;

6) when $d > R_s$ and $d \le R_s + R_p$.

An example of calculating the blocked area is as follow.

In condition 2 shown in **Figure 4**, the blocked area = area 1 + area 2. We can calculate the sectorial area of the star and the planet first. And then we can calculate area1 by subtract the sectorial area of the star with triangle OAB, and use the same method to calculate area 2.

We can use some similar methods to calculate the areas in other conditions.

Using different methods in these conditions, I get the light curve of the star (see **Figure 5**).

2.5. Using the Data of Kepler-19

I use the data of Kepler-19 of Wikipedia (shown in **Table 3**) to run my program assuming that all the orbits of the Kepler-19 system are in the same plane. Following are my results (see **Figures 6-8**).

Table 3. The data of Kepler-19 system.

Name	Mass (kg)	Eccentricity	The average period (day)	Radius (m)
Kepler-19	0.936 × mass of the sun	0	0	0.85 × radius of the sun
Kepler-19b	8.4 × mass of the earth	0.12	9.28716	2.209 × radius of the earth
Kepler-19c	13.1 × mass of the earth	0.21	28.731	unknown
Kepler-19d	22.5 × mass of the earth	0.05	62.95	unknown

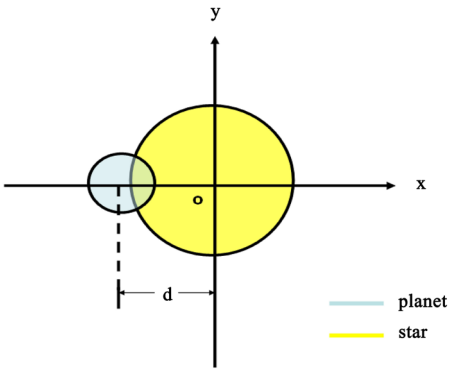


Figure 3. A transiting planet.

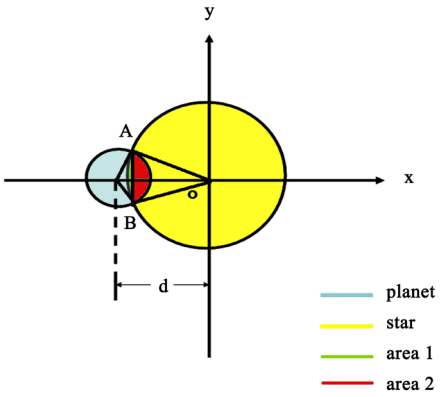


Figure 4. The way to calculate blocked area in condition 2.

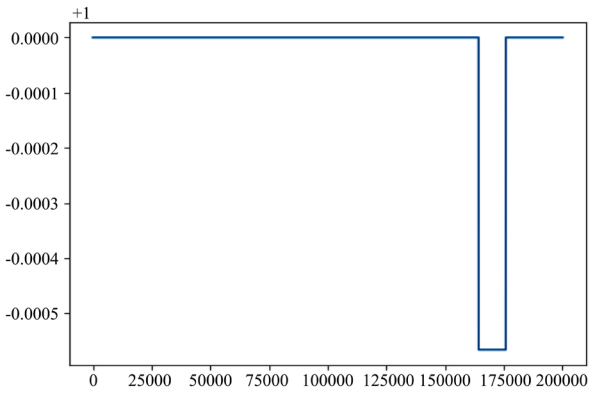


Figure 5. The change in brightness of the star. *x* axis is time in seconds and *y* axis is the ratio of the brightness of the star.

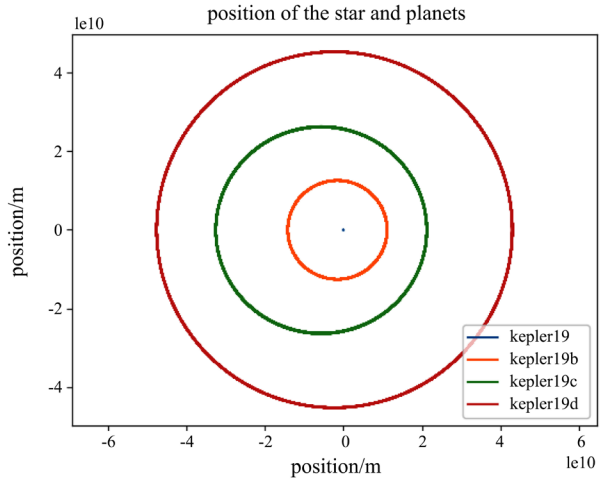


Figure 6. The position of the celestial bodies in Kepler-19 system.

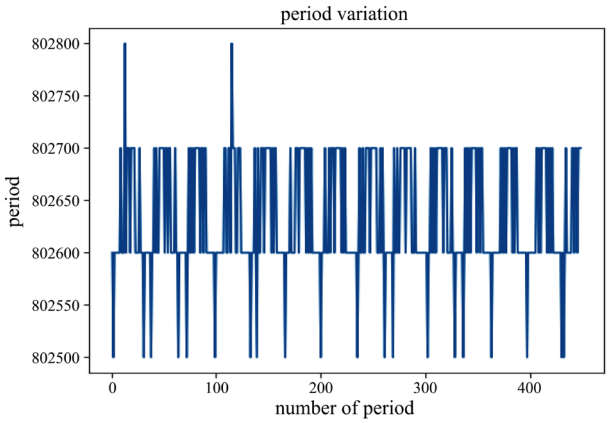


Figure 7. The variation of period.

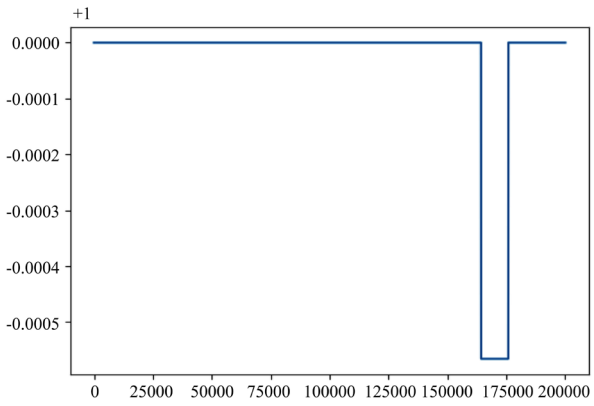


Figure 8. The change in brightness of Kepler-19. *x* axis is time in seconds and *y* axis is the ratio of the brightness of the star.

3. Conclusion

I use python to simulate the gravitational effects on celestial bodies due to other celestial bodies, and I get the orbit of the planets orbiting around Kepler-19 and the light curve of Kepler-19. However, giving the limitation of my laptop, the

time step I use is not short enough to get accurate simulations. Using the time step of 100 seconds and a total time of 360,000,000 seconds, it takes 1 hour and 20 minutes to run the program. To get a more accurate simulation, I need a better computer to run the program. What's more, I will try to do a program that can analyze the light curve of a star to get the information about the mass, radius, and many other properties of the planets which transits can be observed by us and of all the planets that are orbiting around the star.

Conflicts of Interest

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

References

[1] Ballard, S., Fabrycky, D., Fressin, F., Charbonneau, D., Desert, J.-M., Torres, G., Marcy, G., Burke, C.J., Isaacson, H., Henze, C., Steffen, J.H., Ciardi, D.R., Howell, S.B., Cochran, W.D., Endl, M., Bryson, S.T., Rowe, J.F., Holman, M.J., Lissauer, J.J., Jenkins, J.M., Still, M., Ford, E.B., Christiansen, J.L., Middour, C.K., Haas, M.R., Li, J., Hall, J.R., McCauliff, S., Batalha, N.M., Koch, D.G. and Borucki, W.J. (2011) The Kepler-19 System: A Transiting 2.2 $R_{\oplus}$ Planet and a Second Planet Detected via Transit Timing Variations. *The Astrophysical Journal*, **743**, 200. https://doi.org/10.1088/0004-637X/743/2/200

Scientific Research Publishing

Journal of Applied Mathematics and Physics, 2019, 7, 1870-1876
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Diffraction Pattern of a Rotated Grating

Pirooz Mohazzabi, Daniel J. Mattson, Joel Ambriz Ponce Jr.

Department of Mathematics and Physics, University of Wisconsin-Parkside, Kenosha, USA
Email: mohazzab@uwp.edu

How to cite this paper: Mohazzabi, P., Mattson, D.J. and Ponce Jr., J.A. (2019) Diffraction Pattern of a Rotated Grating. *Journal of Applied Mathematics and Physics*, **7**, 1870-1876.
https://doi.org/10.4236/jamp.2019.78128

Received: July 14, 2019
Accepted: August 20, 2019
Published: August 23, 2019

Abstract

Diffraction patterns of a rotated grating are investigated from both theoretical and experimental points of view. It is shown that as the grating rotates, the angle of deviation of each diffracted line relative to the incident light goes through a minimum, and that the angle between any two successive diffracted lines goes through a minimum value which is the same for all adjacent diffracted lines. It is also shown that the angle between diffraction lines with $n=1$ and $n=-1$ is not sensitive to small variations of the grating from being normal to the incident light. Finally, a method is suggested for determining the line distance of a diffraction grating with high accuracy.

Keywords

Diffraction Pattern, Rotated Grating, Angle of Deviation

1. Introduction and Theory

Rotating diffraction gratings have a number of applications. For example, a pair of counter rotating diffraction gratings can be used in laser beam scanners for obtaining a linear scan rate for a flat field recorder [1]. The Doppler frequency shift in various diffraction orders produced as a result of a rotating radial diffraction grating can be used for optical modulation [2]. Rotating all-glass diffraction gratings can be used as beam splitting frequency shifter in laser Doppler anemometers [3]. In this article, we study the general behavior of various diffracted lines as a result of rotation of the diffraction grating.

When a diffraction grating of line distance d is exposed to a light ray of wavelength λ, the general equation for the condition of constructive interference is given by [4]

$$d\left(\sin\beta-\sin\alpha\right)=n\lambda \tag{1}$$

where n is any integer, β is the angle of nth-order diffraction, and α is the angle of incidence. Both α and β are considered positive if measured coun-

terclockwise from the axis (normal to the diffraction grating) and negative if measured clockwise, as shown in Figure 1.

Now suppose that we illuminate a diffraction grating by a monochromatic ray of light of wavelength λ and then turn the grating, keeping the direction of the incident light fixed, starting from normal incidence, the angle through which the grating turns would become the angle of incidence for light, which is α in Equation (1). However, instead of the diffraction angle β, let us study the angle of deviation θ between a diffracted beam and the incident beam, as shown in Figure 2. As in the case of α and β, the angle θ is positive if measured counterclockwise from the direction of incident light and negative if measured clockwise. Obviously, in terms of the deviation angle θ, Equation (1) can be written as

$$d\left[\sin(\theta+\alpha)-\sin\alpha\right]=n\lambda,\quad n=0,\pm1,\pm2,\cdots \tag{2}$$

where α is negative if the grating turns counterclockwise (Figure 2(a)) and positive if the grating turns clockwise (Figure 2(b)).

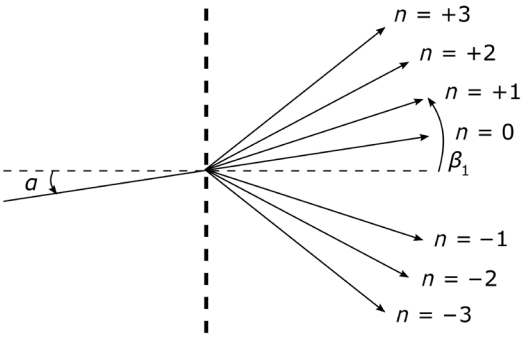
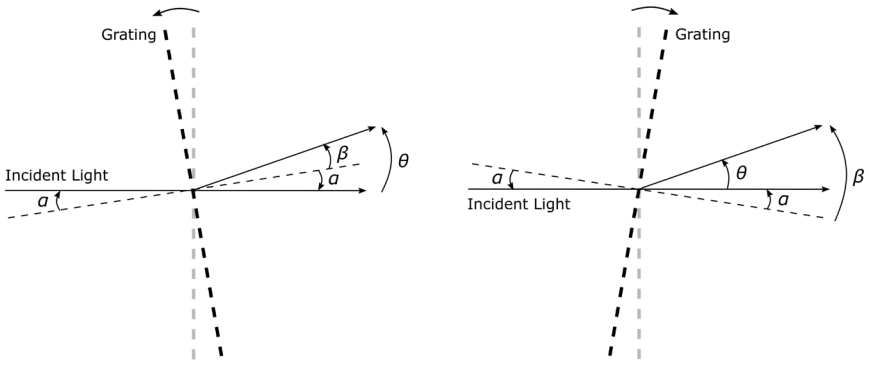


Figure 1. Slant incidence of a monochromatic light on a diffraction grating. For the situation shown α and β_1 are both positive.



(a) (b)

Figure 2. A diagram showing the angle of deviation θ between a diffracted ray of light and the incident ray. The diffraction grating is rotated (a) counterclockwise ($\alpha<0$) and (b) clockwise ($\alpha>0$) relative to normal incidence.

The case $n=0$ is trivial and leads to $\theta=0$. We are, however, interested in the higher-order diffractions and the behavior of θ as a function of α. We shall only consider counterclockwise rotations of the diffraction grating ($\alpha<0$) since clockwise rotations can be obtained from the former by a symmetry operation on the experimental setup. From Equation (2) we have

$$\theta=\sin^{-1}\left(\frac{n\lambda}{d}+\sin\alpha\right)-\alpha \tag{3}$$

Let us investigate the behavior of this function, in particular, find out if there are any maxima or minima in θ,

$$\frac{\mathrm{d}\theta}{\mathrm{d}\alpha}=\frac{\cos\alpha}{\sqrt{1-\left(\frac{n\lambda}{d}+\sin\alpha\right)^2}}-1 \tag{4}$$

This expression becomes zero when

$$\sqrt{1-\left(\frac{n\lambda}{d}+\sin\alpha\right)^2}=\cos\alpha \tag{5}$$

which reduces to

$$\sin\alpha=-\frac{n\lambda}{2d} \tag{6}$$

and becomes infinity if

$$\sqrt{1-\left(\frac{n\lambda}{d}+\sin\alpha\right)^2}=0 \tag{7}$$

which reduces to

$$\frac{n\lambda}{d}+\sin\alpha=\pm1 \tag{8}$$

In the last equation, the positive root will not be considered since it gives $\alpha>0$ (up to $n=5$ for the choice of $\lambda=632.8\ \text{nm}$ and $d=3342\ \text{nm}$), corresponding to a clockwise rotation. Therefore, the possibility of a minimum or maximum in Equation (3) exist only if one of the following conditions is satisfied:

$$\sin\alpha=-\frac{n\lambda}{2d}\quad\text{or}\quad\sin\alpha=-1-\frac{n\lambda}{d} \tag{9}$$

We now consider each case separately.

Case 1: $\sin\alpha=-\dfrac{n\lambda}{2d}$

For the counterclockwise rotations ($\alpha<0$) that we are considering, we must have $n>0$. Furthermore, substituting the above equation into Equation (2), we obtain

$$\sin(\theta+\alpha)=\frac{n\lambda}{2d} \tag{10}$$

Then adding the first of Equations (9) and (10) gives

$$\sin(\theta+\alpha)+\sin\alpha=0 \tag{11}$$

or

$$2\sin\left(\frac{\theta+2\alpha}{2}\right)\cos\left(\frac{\theta}{2}\right)=0 \tag{12}$$

which holds only if $\theta=-2\alpha$. It is easily verified that under this condition, $\mathrm{d}^2\theta/\mathrm{d}\alpha^2>0$, and hence θ is a minimum.

Case 2: $\sin\alpha=-1-\dfrac{n\lambda}{d}$

Clearly this can hold only if $n<0$ which corresponds to the diffracted rays below the normal to the grating (Figure 1). Substituting this result in Equation (1), we find $\sin\beta=-1$ or $\beta=-\pi/2$, which means that the diffracted ray is in the plane of the grating.

Therefore, our theoretical analysis shows that when the diffraction grating turns counterclockwise the diffracted rays with $n>0$ go through a minimum angle of deviation. The minimum angle of deviation for the *n*th-order diffracted beam takes place when the grating rotates through an angle α_n relative to the normal incidence, given by

$$\alpha_n=-\sin^{-1}\left(\frac{n\lambda}{2d}\right) \tag{13}$$

The minimum angle of deviation for the *n*th-order diffraction is then given by

$$\theta_n(\min)=-2\alpha_n=2\sin^{-1}\left(\frac{n\lambda}{2d}\right) \tag{14}$$

There are no other minima or maxima in the angle of deviation.

2. Experiment and Results

In order to test these results, we used a diffraction grating with a line distance of $d=3342\pm10$ nm, and a 0.95 mW He-Ne laser ($\lambda=632.8$ nm). The line distance of the grating was obtained by measuring it at ten different points of the grating, using the laser light.

Table 1 shows the results of our measurements and their comparison with the calculated values. Each measured value reported in the table is the average of ten runs along with its standard deviation. The standard deviations of the theoretical values are calculated using propagation of errors and the standard deviation in *d*. As can be seen from Table 1, the agreement between the measured values and the calculated values is excellent.

Another interesting feature of these diffraction patterns is that the angle through which the grating has to turn for a diffraction line to go to its minimum deviation increases with the order of diffraction *n*. As a result, the angle between any two adjacent diffracted rays decreases first and then increases as the grating turns, resulting in a minimum value. To see this, we write Equation (3) for *n* and again for $n+1$, and subtract them to get

Table 1. Calculated and measured values of the rotation angle of the grating ($-\alpha_n$) at which the angle of deviation becomes a minimum, $\theta_n(\min)$. All angles are in degree.

	$-\alpha_n$		$\theta_n(\min)$	
n	calc	meas	calc	meas
1	5.43 ± 0.02	5.5 ± 0.1	10.87 ± 0.03	10.93 ± 0.02
2	10.91 ± 0.03	11.1 ± 0.2	21.83 ± 0.07	21.89 ± 0.04
3	16.50 ± 0.05	16.6 ± 0.3	33.00 ± 0.10	33.14 ± 0.03
4	22.25 ± 0.07	22.5 ± 0.2	44.51 ± 0.14	44.68 ± 0.02
5	28.25 ± 0.09	28.6 ± 0.2	56.51 ± 0.18	56.75 ± 0.02

$$\theta_{n+1}-\theta_n=\sin^{-1}\left[\frac{(n+1)\lambda}{d}+\sin\alpha\right]-\sin^{-1}\left[\frac{n\lambda}{d}+\sin\alpha\right] \tag{15}$$

Then differentiation of this equation with respect to α gives

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}(\theta_{n+1}-\theta_n)=\frac{\cos\alpha}{\sqrt{1-\left[\frac{(n+1)\lambda}{d}+\sin\alpha\right]^2}}-\frac{\cos\alpha}{\sqrt{1-\left[\frac{n\lambda}{d}+\sin\alpha\right]^2}} \tag{16}$$

which vanishes if and only if

$$\alpha=-\sin^{-1}\left[\frac{\lambda}{d}\left(n+\frac{1}{2}\right)\right] \tag{17}$$

It can easily be verified that when this condition is met, $\theta_{n+1}-\theta_n$ is a minimum. Therefore, when α satisfies Equation (17), the angle between diffraction lines of order n and $n+1$ becomes a minimum with a value given by

$$(\theta_{n+1}-\theta_n)_{\min}=2\sin^{-1}\left(\frac{\lambda}{2d}\right) \tag{18}$$

which is independent of n. Therefore the minimum angle reached between any two adjacent diffraction lines is the same.

Table 2 shows the values of α and $(\theta_{n+1}-\theta_n)_{\min}$ calculated from Equations (17) and (18), respectively. In this case it was difficult to find the value of α for which $(\theta_{n+1}-\theta_n)$ goes to a minimum by watching the diffracted beams, as both θ_n and θ_{n+1} changed with α. We, therefore. measured θ_n and θ_{n+1} each as a function of α and then plotted $\theta_{n+1}-\theta_n$ as a function of α. The minimum of each graph did in fact take place at a value of α very close to that predicted by Equation (17). The measured values of $(\theta_{n+1}-\theta_n)_{\min}$ are given in the last column of Table 2.

It should be pointed out that some of the results discussed in this work were experimentally observed by Lock [5]. However, as explained in the next section, he only provided a limited semi-quantitative analysis for his observations. Furthermore, the common practice in the diffraction grating measurements where the angle of diffraction is measured on both sides of the central maximum and then the average is calculated, is not falsified by the above asymmetric behavior

Table 2. Calculated and measured values of the minimum angle between adjacent diffraction line, and the rotation angles of the grating at which they occur. All angles are in degree.

	$(\theta_{n+1}-\theta_n)_{\min}$	
$-\alpha$	calc	meas
5.4	10.87	10.90
16.5	10.87	10.87
28.3	10.87	10.88
41.5	10.87	10.93
58.4	10.87	10.88

of the diffraction angles. In fact, it turns out that the angle between the diffraction maxima with $n=1$ and $n=-1$ is not sensitive to small deviations of the grating from being normal to the incident beam. To see this we consider the angle between the diffraction maxima with n and $-n$:

$$\theta_n-\theta_{-n}=\sin^{-1}\left(\frac{n\lambda}{d}+\sin\alpha\right)-\sin^{-1}\left(\frac{-n\lambda}{d}+\sin\alpha\right) \tag{19}$$

Taking the derivative of this angle with respect to α, we find

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}(\theta_n-\theta_{-n})=\frac{\cos\alpha}{\sqrt{1-\left(\frac{n\lambda}{d}+\sin\alpha\right)^2}}-\frac{\cos\alpha}{\sqrt{1-\left(\frac{n\lambda}{d}-\sin\alpha\right)^2}} \tag{20}$$

which vanishes if and only if $\alpha=0$. Therefore, the angle $\theta_n-\theta_{-n}$ having a stationary value at $\alpha=0$, is not affected appreciably by small deviations of α from zero.

Finally, as an application of the above results, we rewrite Equation (14) as

$$d=\frac{n\lambda}{2\sin\left[\frac{\theta_n(\min)}{2}\right]} \tag{21}$$

Since θ_n as a function of α has a very flat minimum, $\theta_n(\min)$ can be measured very accurately for a grating and hence the line distance d can be obtained from Equations (21) with high accuracy. Furthermore, in a single experiment data can be collected on several diffraction orders, resulting in a statistically even more accurate measurement of d. In these measurements the task of making the diffraction grating perpendicular to the incident light, which is normally required in the usual diffraction experiments, is eliminated.

3. Summary

We investigated the diffraction patters of a rotated grating, from both theoretical and experimental points of view. The results of this investigation show several interesting behaviors of the system. First, as the grating rotates, the angle of deviation of each diffracted line relative to the incident light goes through a mini-

mum. Although this behavior has been reported by Lock [5], he only provided a semi-quantitative analysis of the problem based on two competing effects. Furthermore, Lock only considered the first- and second-order diffraction lines. In this work, we have provided a complete quantitative analysis of the problem and have considered up to the fifth-order diffraction.

Second, we have shown that the rotation angle of the grating that results in a minimum angle of deviation for a given diffracted line increases with the order of diffraction. As a result, the angle between any two successive diffracted lines first decreases and then increases, resulting in a minimum value which is the same for any two adjacent diffracted lines. This result has not been reported previously.

Third, it is shown that the angle between diffraction lines with $n=1$ and $n=-1$ is not sensitive to small variations of the grating from being normal to the incident light, which justifies the commonly practiced procedure in diffraction experiments, again, an observation that is missing from previous investigations.

Finally, we have suggested a new method for determining the line distance of a diffraction grating with high accuracy.

Acknowledgements

This work was supported in part by a URAP grant from the University of Wisconsin-Parkside.

Conflicts of Interest

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

References

[1] Wyant, J.C. (1975) Rotating Diffraction Grating Laser Beam Scanner. *Applied Optics*, **14**, 1057-1058. https://doi.org/10.1364/AO.14.001057

[2] Stevenson, W.H. (1970) Optical Frequency Shifting by Means of a Rotating Diffraction Grating. *Applied Optics*, **9**, 649-652. https://doi.org/10.1364/AO.9.000649

[3] Oldengarm, J. (1977) Development of Rotating Diffraction Gratings and Their Use in Laser Anemometry. *Optics and Laser Technology*, **4**, 69-71. https://doi.org/10.1016/0030-3992(77)90024-X

[4] Hecht, E. and Zajac, A. (1979) Optics. Addison-Wesley, Menlo Park, 357.

[5] Lock, J.A. (1985) The Rotated Diffraction Grating: A Laboratory Experiment. *The Physics Teacher*, **23**, 226. https://doi.org/10.1119/1.2341789

Journal of Applied Mathematics and Physics, 2019, 7, 1877-1892
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Some New Applications of Elzaki Transform for Solution of Linear Volterra Type Integral Equations

Saad Sharjeel, Mushtaq Ahmed Khan Barakzai

College of Computer Science and Information System (CCSIS), Institute of Business Management (IoBM), Karachi, Pakistan
Email: std_15459@iobm.edu.pk, mushtaq.barakzai@iobm.edu.pk

How to cite this paper: Sharjeel, S. and Barakzai, M.A.K. (2019) Some New Applications of Elzaki Transform for Solution of Linear Volterra Type Integral Equations. *Journal of Applied Mathematics and Physics*, **7**, 1877-1892.
https://doi.org/10.4236/jamp.2019.78129

Received: June 29, 2019
Accepted: August 20, 2019
Published: August 23, 2019

Abstract

Volterra type integral equations have diverse applications in scientific and other fields. Modelling physical phenomena by employing integral equations is not a new concept. Similarly, extensive research is underway to find accurate and efficient solution methods for integral equations. Some of noteworthy methods include Adomian Decomposition Method (ADM), Variational Iteration Method (VIM), Method of Successive Approximation (MSA), Galerkin method, Laplace transform method, etc. This research is focused on demonstrating Elzaki transform application for solution of linear Volterra integral equations which include convolution type equations as well as one system of equations. The selected problems are available in literature and have been solved using various analytical, semi-analytical and numerical techniques. Results obtained after application of Elzaki transform have been compared with solutions obtained through other prominent semi-analytic methods *i.e.* ADM and MSA (limited to first four iterations). The results substantiate that Elzaki transform method is not only a compatible alternate approach to other analytic methods like Laplace transform method but also simple in application once compared with methods ADM and MSA.

Keywords

Elzaki Transform, Linear Volterra Integral Equation, Adomian Decomposition Method, Method of Successive Approximation, Laplace Elzaki Duality (LED)

1. Introduction

Integral equations find their application in physical sciences, finance, etc. Diffraction problems, water waves, scattering in quantum mechanics are often

modelled using integral equations [1]. The unknown function which is required to be determined is denoted by $u(x)$ and it appears under the sign of integration. $K(x, t)$ is called "kernel" function whereas $g(x)$ and $h(x)$ are the limits of integration. [1] has provided excellent introduction on Integral equations along with solution methods. General form of an Integral equation can be given as under:

$$u(x) = f(x) + \lambda \int_{h(x)}^{g(x)} k(x,t)u(t)\mathrm{d}t \tag{1}$$

An Integro-differential equation is a bit different from Integral equation as it contains unknown function $u(x)$ which appears under integral sign and also has ordinary derivative of unknown function. For the Integro-differential equation, general form can be given as:

$$u^n(x) = f(x) + \lambda \int_{h(x)}^{g(x)} k(x,t)u(t)\mathrm{d}t \tag{2}$$

The system of integral or Integro-differential equations has two or more equations with two or more variables which are required to be determined. Volterra integral, integro-differential as well as related system of equations contains at least one variable limit of integration. Extensive mathematical techniques are available for finding analytic (exact), approximate analytic as well as numerical solutions of integral equations. Some of the techniques include Adomian Decomposition Method (ADM), Homotopy Perturbation Method (HPM), Homotopy Analysis Method (HAM), Variational Iteration Method (VIE), etc. Transformation methods like Sumudu transform, Laplace transform, etc. are also being used to find solution of integral equations of various types and classes. However, new methods are continuously being explored throughout the world. Elzaki transform has been introduced by [2] and there is a growing interest of researchers in finding various applications of said transformation method. A large number of mathematical problems have been solved using Elzaki transform method. However, with respect to finding its new applications, there is great potential available. Volterra type integral equations have been earlier solved using numerous methods available by mathematicians. In this paper some new applications of Elzaki transform have been discussed to find analytic solution of linear Volterra type integral equations which include convolution type as well as system of equations. The problems selected for demonstrating Elzaki transform application are those which have not been solved earlier using said transformation method. The analytic solutions obtained after application of Elzaki transform have been compared with results obtained through famous semi-analytical methods *i.e.* ADM and MSA (which have been restricted to first four iterations). The results establish the accuracy and simplicity of Elzaki transform method and also attest its compatibility with Laplace transform.

2. Literature Review

Integro-differential equations with bulge function have been examined by [3]. Numerical solution has been obtained by applying trapezoidal rule. For finding

the exact solution; Elzaki and inverse Elzaki transform as well as Taylor series expansion and convolution theorem have been used. Through examples, it has been shown that the approximate solutions acquired by trapezoidal rule are in good agreement with exact solutions obtained through transformation method. [4] conducted comparative study between Adomian Decomposition Method (ADM) and Elzaki transform. Both methods have been used to solve linear partial differential equations with constant coefficients. Elzaki transform method has been used by [5] for solution of systems of linear Integro-differential equations with constant coefficients. Fundamental properties of Elzaki transform have been discussed by [6] and Elzaki transform for comprehensive list of functions has been provided. Furthermore, more general shift theorems have been introduced. Laplace-Elzaki Duality (LED) invoked a complex inverse Elzaki transform, as a Bromwich contour integral formula. [7] researched practical formulae for differentiation of integral transforms used for differential equations with variable coefficients. The transforms which have been checked are Laplace, Sumudu and Elzaki. Moreover, it has been argued that proposed formulae can be applied to almost every equation. [8] proposed shifting theorems for the Elzaki transform to solve initial value problems arising in control engineering. The proposed theorems are composed of u-shifting theorem and time shifting, and the proof is compared with established ones. [9] provided Romberg method for solution of linear Volterra integral equations of second kind. The examples presented in paper show superiority of Romberg method over quadrature method. [10] solved Volterra integral equations with separable kernels using the differential transform method. Approximate solution has been calculated in form of a series with easily calculable terms. Exact solutions of linear as well as nonlinear integral equations have been presented. Results exemplify the reliability of the differential transform method. Collocation method has been presented by [11] for linear Volterra integral equation of the second kind by using Sinc basis functions. Approximate solutions are provided and auxiliary basis functions satisfy four different boundary conditions. Numerical results have been included to confirm efficiency and accuracy of method. [12] solved Volterra integral equations of second kind (convolution type) by using the Elzaki transform. Solution of integro-differential equations using Elzaki transform has been discussed by [13]. Numerical solution of a system of two first order Volterra integro-differential equations arising in ultimate ruin theory has been discussed by [14]. Existence and Uniqueness of Solution of Volterra Integral Equations has been studied by [15]. [16] has used Finite Difference Method for Smooth Solution of Linear Volterra Integral Equations. Numerical Solutions of Volterra Integral Equations Using Galerkin method with Hermite Polynomials have been discussed by [17].

3. Mathematical Foundations of Elzaki Transform

Elzaki transform has been derived from classical Fourier Integral.

By definition, Elzaki transform is defined for given set A as:

$$A=\left\{ f(t)\,|\,\exists M,k_1,k_2>0\,|\,f(t)<Me^{\frac{-t}{k_j}},\text{ if } t\in(-1)^j\times[0,\infty)\right\} \tag{3}$$

For the given function M should be finite however k_1 and k_2 may be finite or infinite.

Elzaki transform is denoted by $E(.)$ and is given as

$$E\left[f(t)\right]=T(p)=p\int_0^{\infty} f(t)e^{-t/p}dt, t\geq 0,\ \ k_1\leq p\leq k_2 \tag{4}$$

here $T(p)$ is Elzaki transform of integral function $f(t)$.

For $f(t)$ we assume that the integral given in Equation (4) exists.

Let $f(t)=1$

$$E[1]=T(p)=p\int_0^{\infty}1\cdot e^{-t/p}dt=p^2 \tag{5}$$

Let $f(t)=t$

$$E[t]=T(p)=p\int_0^{\infty}t\cdot e^{-\frac{t}{p}}dt=p^3 \tag{6}$$

For the nth order

$$E\left[t^n\right]=n!\,p^{n+2} \tag{7}$$

3.1. Elzaki Transform of Common Functions

Elzaki transform of some common functions is given as under Elzaki transform of exponential function

$$E\left(e^{at}\right)=T(p)=p\int_0^{\infty}e^{at}\cdot e^{-\frac{t}{p}}dt=\frac{p^2}{1-ap} \tag{8}$$

Elzaki transform of Sin function

$$E\left(\sin(at)\right)=\frac{ap^3}{a^2p^2+1} \tag{9}$$

Elzaki transform of Cosine function

$$E\left(\cos(at)\right)=\frac{p^2}{a^2p^2+1} \tag{10}$$

Elzaki transform of Sin hyperbolic function

$$E\left(\sinh(at)\right)=\frac{ap^3}{1-a^2p^2} \tag{11}$$

Similarly Elzaki transform for derivatives of a function can be given as

$$E\left(f'(t)\right)=\frac{T(p)}{p}-vf(0) \tag{12}$$

$$E\left(f''(t)\right)=\frac{T(p)}{p^2}-f(0)-vf'(0) \tag{13}$$

$$E\left(f^n(t)\right)=\frac{T(p)}{p^n}-\sum_{k=0}^{n-1}p^{2-n+k}f^k(0) \tag{14}$$

3.2. Laplace Elzaki Duality (LED)

For function $f(t)$ which belongs to set A,

$$f(t)\in A=\left\{f(t)\,|\,\exists M,k_1,k_2>0,\text{ such that }|f(t)|<M\mathrm{e}^{\frac{|t|}{k_j}},\text{ if } t\in(-1)^j\times[0,\infty)\right\} \tag{15}$$

Since Laplace transform is given as

$$F(s)=L\left(f(t)\right)=\int_0^\infty \mathrm{e}^{-st}f(t)\mathrm{d}t \tag{16}$$

Inter-conversion between Laplace and Elzaki transform can be given by

$$T(v)=vF\left(\frac{1}{v}\right) \tag{17}$$

$$F(s)=sT\left(\frac{1}{s}\right) \tag{18}$$

4. Examples

Example 4.1

$$u(x)=4x+2x^2-\int_0^x u(t)\mathrm{d}t \tag{19}$$

This problem has been taken from [18].

Taking Elzaki transform on both sides of Equation (19)

$$E\left[u(x)\right]=4\left[E(x)\right]+2\left[E\left(x^2\right)\right]-E\left[\int_0^x u(t)\mathrm{d}t\right] \tag{20}$$

$$T[v]=4v^3+2\left(2!v^{2+2}\right)-vT[v] \tag{21}$$

$$T[v]=4v^3+4v^4-vT[v] \tag{22}$$

$$T[v]+vT[v]=4\left(v^3+v^4\right) \tag{23}$$

$$T[v]=\frac{4v^3(1+v)}{1+v} \tag{24}$$

After simplification,

$$T[v]=4v^3 \tag{25}$$

Taking Inverse Elzaki transform on both sides of Equation (25), we get,

$$u(x)=4x \tag{26}$$

Hence, required analytic solution is obtained.

Example 4.2

$$u(x)=x^2+\frac{1}{12}x^4+\int_0^x\left((t-x)u(t)\right)\mathrm{d}t \tag{27}$$

This problem has been discussed by [18].

Taking derivatives and applying Leibniz rule to equation Equation (27)

$$u'(x) = 2x + \frac{1}{3}x^3 - \int_0^x u(t)\,\mathrm{d}t \tag{28}$$

$$u''(x) = 2 + x^2 - u(x) \tag{29}$$

$$u''(x) + u(x) = 2 + x^2, \text{ with } u(0) = 0, u'(0) = 0 \tag{30}$$

Since,

$$E\left[u''(x)\right] = \frac{T[v]}{v^2} - u(0) - vu'(0) \tag{31}$$

After plugging initial values we obtain,

$$E\left[u''(x)\right] = \frac{T[v]}{v^2} \tag{32}$$

Hence, the Elzaki transform of given equation will be

$$E\left[u''(x)\right] + E\left[u(x)\right] = E[2] + E\left[x^2\right] \tag{33}$$

Or,

$$\frac{T[v]}{v} + T[v] = 2 + 2v^4 \tag{34}$$

$$T[v] + v^2 T[v] = 2v^2\left(v^2 + v^4\right) \tag{35}$$

$$T[v] = \frac{2v^2\left(v^2 + v^4\right)}{1 + v^2} \tag{36}$$

Taking inverse Elzaki transform on both sides of Equation (36) we get,

$$u(x) = E^{-1}\left[\frac{2v^2\left(v^2 + v^4\right)}{1 + v^2}\right] \tag{37}$$

$$u(x) = E^{-1}\left[2v^4\right] \tag{38}$$

$$u(x) = x^2 \tag{39}$$

Hence, required analytic solution is obtained.

Example 4.3

Consider Volterra Integral equation of Convolution type:

$$u(x) = 2 + \int_0^x (x - t)u(t)\,\mathrm{d}t \tag{40}$$

This problem has been taken from [18],

Taking Elzaki transform on both sides of Equation (40),

$$E\left[u(x)\right] = E[2] + E\left[\int_0^x \left((x - t)u(t)\right)\mathrm{d}t\right] \tag{41}$$

$$T[v] = 2v^2 + \frac{1}{v}\left[v^3 T[v]\right] \tag{42}$$

$$T[v] = 2v^2 + v^2 T[v] \tag{43}$$

$$T[v]-v^2T[v]=2v^2 \tag{44}$$

$$T[v]=\frac{2v^2}{1-v^2} \tag{45}$$

Taking Inverse Elzaki transform on both sides of Equation (45),

$$u(x)=2\cosh x \tag{46}$$

Hence, required analytic solution is obtained.

Example 4.4

This problem has been solved by [11] using Sinc basis functions,

$$u(x)=x-x^2+\frac{x^3}{6}-\frac{x^4}{12}+\int_0^x\big((t-x)u(t)\big)\mathrm{d}t, 0\le x\le 1 \tag{47}$$

Taking Elzaki transform on both sides of Equation (47),

$$E[u]=E[x]-E\left[x^2\right]+\frac{1}{6}E\left[x^3\right]-\frac{1}{12}E\left[x^4\right]+E\left[\int_0^x\big((t-x)u(t)\big)\mathrm{d}t\right] \tag{48}$$

$$T[v]=v^3-2v^4+v^5-2v^6-\frac{1}{v}\big(E[t]E[u]\big) \tag{49}$$

$$T[v]=v^3-2v^4+v^5-2v^6-v^2T[v] \tag{50}$$

$$T[v]=v^3-2v^4 \tag{51}$$

Taking Inverse Elzaki transform on both sides,

$$u(x)=x-x^2 \tag{52}$$

Hence, required analytic solution is obtained.

Example 4.5

Consider convolution type linear Volterra integral equation. [10] solved the problem using Differential Transform Method (DTM),

$$u(x)=1-x-\frac{x^2}{2}+\int_0^x\big((x-t)u(t)\big)\mathrm{d}t, 0<x<1 \tag{53}$$

Taking Elzaki transform on both sides of Equation (53),

$$E\big[u(x)\big]=E[1]-E[x]-E\left[\frac{x^2}{2}\right]+E\left[\int_0^x\big((x-t)u(t)\big)\mathrm{d}t\right] \tag{54}$$

$$T[v]=v^2-v^3-v^4+\frac{1}{v}\big[E[t]E[u]\big] \tag{55}$$

$$T[v]=v^2-v^3-v^4+v^2T[v] \tag{56}$$

$$T[v]=\frac{v^2-v^3-v^4}{1-v^2} \tag{57}$$

Through Taylor expansion about $v=0$,

$$T[v]=v^2-v^3-v^5-v^7-v^9-\cdots \tag{58}$$

Taking Inverse Elzaki transform on both sides of Equation (58),

$$E^{-1}\left[T\left[v\right]\right]=E^{-1}\left[v^2\right]-E^{-1}\left[v^3\right]-E^{-1}\left[v^5\right]-E^{-1}\left[v^7\right]-\cdots \tag{59}$$

In series form,

$$u\left(x\right)=1-\left(x+\frac{x^3}{6}+\frac{x^5}{120}+\frac{x^7}{5040}+\frac{x^9}{362880}+O\left(x^{11}\right)\right) \tag{60}$$

In closed form,

$$u\left(x\right)=1-\sinh\left(x\right) \tag{61}$$

Similarly we can also use LED and demonstrate that it gives same result as obtained using Taylor expansion. Since,

$$F\left(s\right)=sT\left(\frac{1}{s}\right)$$

We can write Equation as,

$$F\left(s\right)=\mathcal{L}_s^{-1}\left[\frac{s\left(-\left(\frac{1}{s}\right)^4-\left(\frac{1}{s}\right)^3+\left(\frac{1}{s}\right)^2\right)}{1-\left(\frac{1}{s}\right)^2}\right]\left(x\right) \tag{62}$$

$$u\left(x\right)=\frac{\mathrm{e}^{-x}}{2}-\frac{\mathrm{e}^{x}}{2}+1 \tag{63}$$

Writing in closed form,

$$u\left(x\right)=1-\sinh\left(x\right)$$

Hence, required analytic solution is obtained.

Example 4.6

This system of equations has been solved by [19] using rationalized Haar functions.

$$y_1\left(t\right)-\int_0^t y_2\left(s\right)\mathrm{d}s=1-t^2 \tag{64}$$

$$y_2\left(t\right)-\int_0^t y_1\left(s\right)\mathrm{d}s=t \tag{65}$$

Take Elzaki transform on both sides of Equation (64) we have,

$$E\left[y_1\left(t\right)\right]-E\left[\int_0^t y_2\left(s\right)\mathrm{d}s\right]=E\left[1\right]-E\left[t^2\right] \tag{66}$$

$$\overline{y}_1\left[v\right]-v\overline{y}_2\left[v\right]=v^2-2v^4 \tag{67}$$

Similarly take Elzaki transform of second equation *i.e.* Equation (65),

$$E\left[y_2\left(t\right)\right]-E\left[\int_0^t y_1\left(s\right)\mathrm{d}s\right]=E\left[t\right] \tag{68}$$

$$\overline{y}_2\left[v\right]-v\overline{y}_1\left[v\right]=v^3 \tag{69}$$

After simplification,

$$\overline{y}_2\left[v\right]=2v^3 \tag{70}$$

Take Inverse Elzaki transform of Equation (70) we have,

$$E^{-1}\left[\bar{y}_2[v]\right] = E^{-1}\left[2v^3\right] \tag{71}$$

$$y_2(t) = 2t \tag{72}$$

And from Equation (69) we have,

$$\bar{y}_2[v] - v\bar{y}_1[v] = v^3$$

$$\bar{y}_1[v] = v^2 \tag{73}$$

Take Inverse Elzaki transform of Equation (73) we have,

$$E^{-1}\left[\bar{y}_1[v]\right] = E^{-1}\left[v^2\right] \tag{74}$$

$$y_1(t) = 1 \tag{75}$$

Equation (72) and Equation (75) provides analytic solution.

5. Results and Discussions

It has been established that Elzaki transform can be easily implemented to find analytic solution of Volterra type integral equations as discussed through various problems in section 4. In this section we shall provide the solution for domain 0 to 1 and for comparison purposes, semi-analytic methods (ADM and MSA with zero as initial guess) solutions (limited to first four iterations), have also been given. It is likely that results of semi-analytic methods will converge to analytic solution if more iteration is carried out; but same will be at cost of increased computational work. All computational work has been carried out in Mathematica version 9.

For Example 4.1 the analytic solution is $u(x) = 4x$ which has also been achieved through application of Elzaki transform while the error for ADM and MSA once restricted to the first four iterations is evident (Table 1). However, ADM in this case produced better results once compared with MSA. Significant deviation of MSA from analytic results is obvious in Figure 1.

For Example 4.2 the analytic solution is $u(x) = x^2$ which has also been achieved through application of Elzaki transform while the error for ADM and MSA once restricted to first four iterations is apparent (Table 2). However, ADM in this case has much better results as compared to MSA. MSA is showing large deviation from analytic results (Figure 2).

For Example 4.3 the analytic solution is $u(x) = 2\cosh x$ which has also been achieved through application of Elzaki transform. ADM has slightly better accuracy once compared with MSA. Table 3 and Figure 3 depict results in numeric and graphical form respectively.

For Example 4.4 the analytic solution is $u(x) = x - x^2$ which has also been achieved through application of Elzaki transform while the error for ADM and MSA once restricted to first four iterations is shown in Table 4. Figure 4 depicts results in graphical form.

Graph plot shows considerable variation of results of MSA from analytic result.

Table 1. Solution of Example 4.1 using different methods and absolute errors for approximate analytic methods.

x	Analytic	ADM	MSA	Abs Error MSA	Abs Error ADM
0	0.000000	0.000000	0.000000	0.000000	0.000000
0.1	0.400000	0.400000	0.400183	0.000183	0.000000
0.2	0.800000	0.799989	0.801600	0.001600	0.000011
0.3	1.200000	1.199919	1.205850	0.005850	0.000081
0.4	1.600000	1.599659	1.614933	0.014933	0.000341
0.5	2.000000	1.998958	2.031250	0.031250	0.001042
0.6	2.400000	2.397408	2.457600	0.057600	0.002592
0.7	2.800000	2.794398	2.897183	0.097183	0.005602
0.8	3.200000	3.189077	3.353600	0.153600	0.010923
0.9	3.600000	3.580317	3.830850	0.230850	0.019683
1	4.000000	3.966667	4.333333	0.333333	0.033333

Table 2. Solution of Example 4.2 using different methods and absolute errors for approximate analytic methods.

x	Analytic	ADM	MSA	Abs Error ADM	Abs Error MSA
0	0.000000	0.000000	0.000000	0.000000	0.000000
0.1	0.010000	0.010000	0.009958	0.000000	0.000042
0.2	0.040000	0.040002	0.039333	0.000002	0.000667
0.3	0.090000	0.090027	0.086625	0.000027	0.003375
0.4	0.160000	0.160146	0.149334	0.000146	0.010666
0.5	0.250000	0.250531	0.223964	0.000531	0.026036
0.6	0.360000	0.361484	0.306023	0.001484	0.053977
0.7	0.490000	0.493435	0.390038	0.003435	0.099962
0.8	0.640000	0.646858	0.469566	0.006858	0.170434
0.9	0.810000	0.822055	0.537223	0.012055	0.272777
1	1.000000	1.018750	0.584722	0.018750	0.415278

Table 3. Solution of Example 4.3 using different methods and absolute errors.

x	Analytic	ADM	MSA	Abs Error ADM	Abs Error MSA
0	2.000000	2.000000	2.000000	0.000000	0.000000
0.1	2.010008	2.010008	2.010008	0.000000	0.000000
0.2	2.040134	2.040134	2.040133	0.000000	0.000000
0.3	2.090677	2.090677	2.090675	0.000000	0.000002
0.4	2.162145	2.162145	2.162133	0.000000	0.000011
0.5	2.255252	2.255252	2.255208	0.000000	0.000044
0.6	2.370930	2.370930	2.370800	0.000001	0.000130
0.7	2.510338	2.510335	2.510008	0.000003	0.000330
0.8	2.674870	2.674862	2.674133	0.000008	0.000737
0.9	2.866173	2.866151	2.864675	0.000022	0.001498
1	3.086161	3.086111	3.083333	0.000050	0.002828

Table 4. Solution of Example 4.4 using different methods and absolute errors.

x	Analytic	ADM	MSA	Abs Error ADM	Abs Error MSA
0	0.000000	0.000000	0.000000	0.000000	0.000000
0.1	0.090000	0.090001	0.089708	0.000001	0.000292
0.2	0.160000	0.160021	0.158000	0.000021	0.002000
0.3	0.210000	0.210147	0.204376	0.000147	0.005624
0.4	0.240000	0.240560	0.229339	0.000560	0.010661
0.5	0.250000	0.251516	0.234402	0.001516	0.015598
0.6	0.240000	0.243278	0.222093	0.003278	0.017907
0.7	0.210000	0.215998	0.195971	0.005998	0.014029
0.8	0.160000	0.169574	0.160641	0.009574	0.000641
0.9	0.090000	0.103489	0.121770	0.013489	0.031770
1	0.000000	0.016667	0.086111	0.016667	0.086111

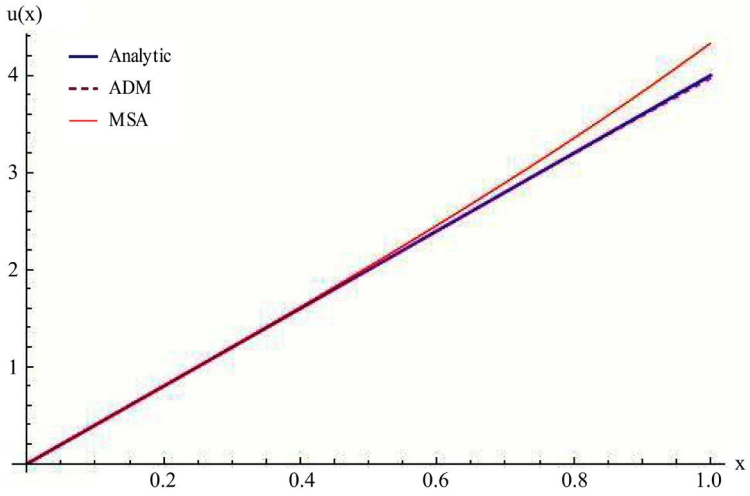


Figure 1. Comparison graph of solution for Example 4.1.

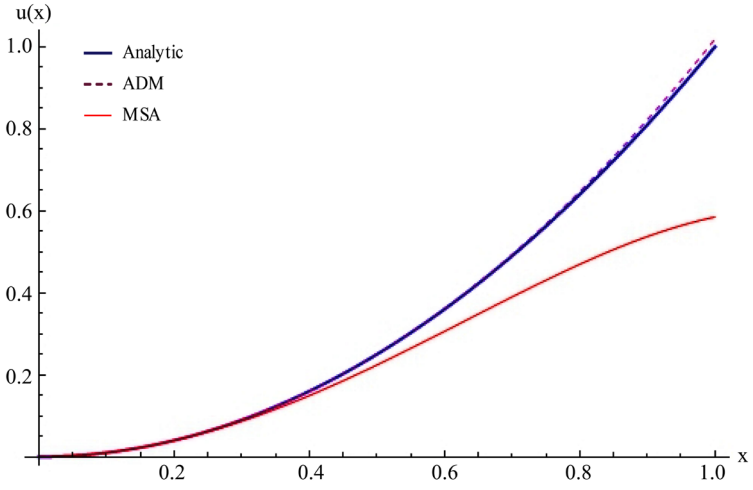


Figure 2. Comparison graph of solution for Example 4.2.

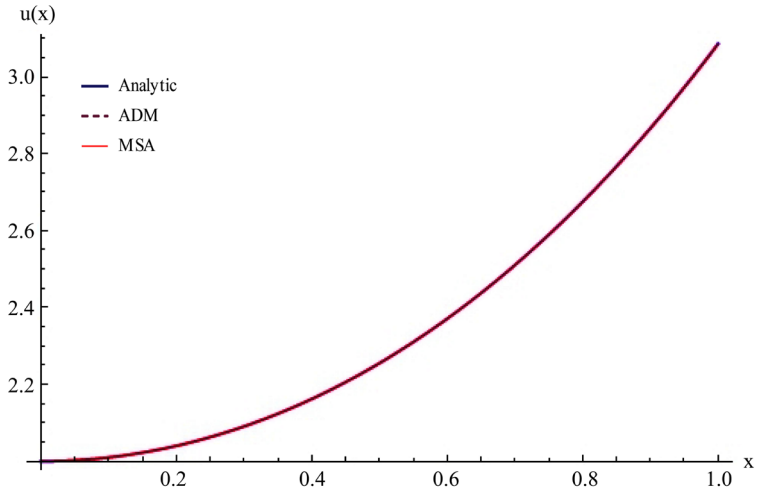


Figure 3. Comparison graph of solution for Example 4.3.

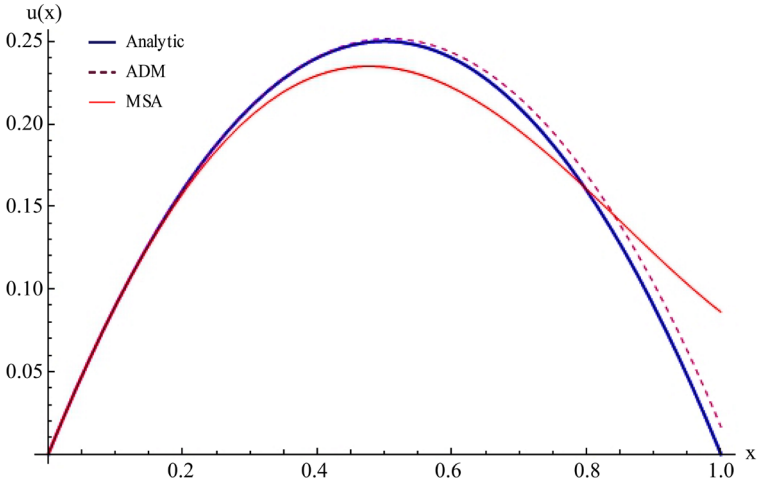


Figure 4. Comparison graph of solution for Example 4.4.

For Example 4.5 the analytic solution is $u(x)=1-\sinh(x)$ which has also been achieved through application of Elzaki transform while the error for ADM and MSA once restricted to first four iterations is evident (Table 5). Figure 5 is showing graphical comparison between methods.

Deviation of MSA solution from analytic result is clearly evident.

For Example 4.6 one analytic solution is $u(t)=y_1(t)=1$ which has also been achieved through application of Elzaki transform. ADM and MSA are showing large deviation from analytic result. Results have been depicted in Table 6(a) while comparison graph between methods has been illustrated at Figure 6(a).

ADM and MSA are showing large deviation from analytic result.

For example 4.6 the second analytic solution is $v(t)=y_2(t)=2t$ which has also been achieved through application of Elzaki transform while the error for ADM and MSA once restricted to first four iterations is obvious. Table 6(b) provides results in numeric form while comparison plot between methods has been given at Figure 6(b).

Table 5. Solution of Example 4.5 using different methods and absolute errors.

x	Analytic	ADM	MSA	Abs Error ADM	Abs Error MSA
0	1.000000	1.000000	1.000000	0.000000	0.000000
0.1	0.899833	0.899833	0.899499	0.000000	0.000334
0.2	0.798664	0.798664	0.795972	0.000000	0.002692
0.3	0.695480	0.695480	0.686282	0.000000	0.009197
0.4	0.589248	0.589248	0.567061	0.000000	0.022186
0.5	0.478905	0.478905	0.434570	0.000000	0.044334
0.6	0.363346	0.363346	0.284548	0.000000	0.078798
0.7	0.241416	0.241415	0.112043	0.000001	0.129373
0.8	0.111894	0.111890	-0.088768	0.000004	0.200662
0.9	-0.026517	-0.026526	-0.324779	0.000010	0.298262
1	-0.175201	-0.175223	-0.604167	0.000022	0.428965

Table 6. (a) Solution of Example 4.6 using different methods and absolute errors for $u(t)=y_1(t)$. (b) Solution of Example 4.6 using different methods and absolute errors for $v(t)=y_2(t)$.

(a)

x	Analytic	ADM	MSA	Abs Error ADM	Abs Error MSA
0	1.000000	1.000000	1.000000	0.000000	0.000000
0.1	1.000000	1.000158	0.999992	0.000158	0.000008
0.2	1.000000	1.001195	0.999867	0.001195	0.000133
0.3	1.000000	1.003785	0.999325	0.003784	0.000675
0.4	1.000000	1.008363	0.997867	0.008363	0.002133
0.5	1.000000	1.015104	0.994792	0.015104	0.005208
0.6	1.000000	1.023904	0.989200	0.023904	0.010800
0.7	1.000000	1.034357	0.979992	0.034357	0.020008
0.8	1.000000	1.045739	0.965867	0.045739	0.034133
0.9	1.000000	1.056984	0.945325	0.056983	0.054675
1	1.000000	1.066667	0.916667	0.066667	0.083333

(b)

x	Analytic	ADM	MSA	Abs Error ADM	Abs Error MSA
0	0.000000	0.000000	0	0.000000	0.000000
0.1	0.200000	0.200000	0.199833	0.000000	0.000167
0.2	0.400000	0.399995	0.398667	0.000005	0.001333
0.3	0.600000	0.599960	0.5955	0.000040	0.004500
0.4	0.800000	0.799829	0.789333	0.000171	0.010667
0.5	1.000000	0.999479	0.979167	0.000521	0.020833
0.6	1.200000	1.198704	1.164	0.001296	0.036000
0.7	1.400000	1.397199	1.342833	0.002801	0.057167
0.8	1.600000	1.594539	1.514667	0.005461	0.085333
0.9	1.800000	1.790159	1.6785	0.009842	0.121500
1	2.000000	1.983333	1.833333	0.016667	0.166667

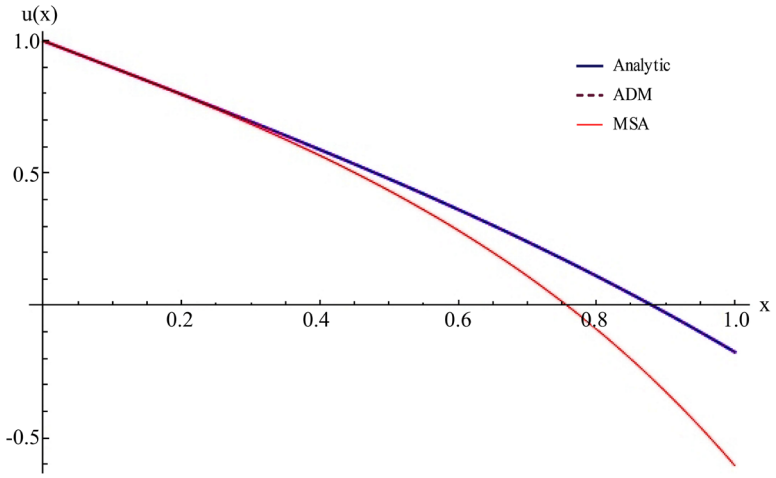


Figure 5. Comparison graph of solution for Example 4.5.

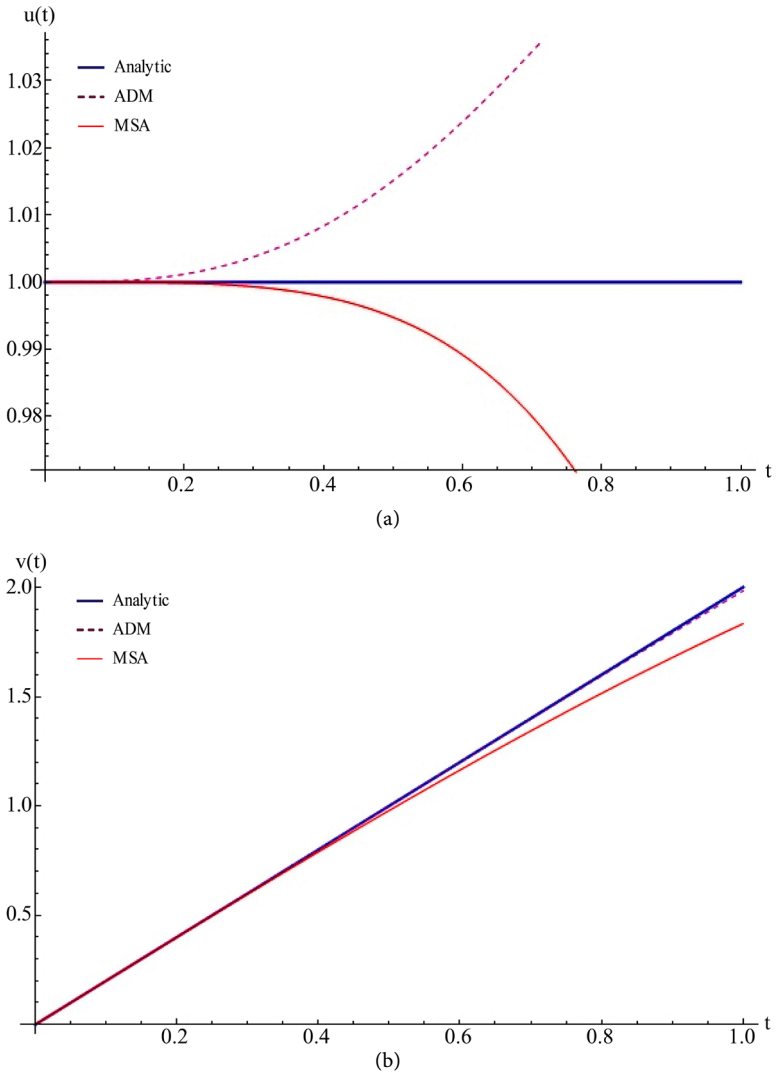


Figure 6. (a). Comparison graph of solution for example 4.6 (for $u(t) = y_1(t)$); (b) Comparison graph of solution for Example 4.6 (for $v(t) = y_2(t)$).

ADM has considerably good accuracy once compared with MSA.

Accurate analytic solution achievement is useful advantage of Elzaki transform. Moreover, the application process is very simple. Methods like ADM despite having extensive applications and advantages, has not been able to produce analytic results for problems once restricted to fewer iterations (in this case four iterations). MSA further produced results of lower accuracy once compared with ADM.

6. Conclusion

In this research Elzaki transform has been successfully applied to linear Volterra type integral equations to find analytic solutions. It has been established that Elzaki transform is a robust compatible alternative to other well-known analytic methods. Moreover, in comparison to notable semi/approximate analytic methods like Adomian Decomposition Method and Method of Successive Approximations, it is not only accurate but often easier to apply. Results presented in research substantiate this claim. The research can be further extended by discussing application of Elzaki transform for linear Volterra type integral equations with separable kernels.

Acknowledgements

The authors are grateful to the Dr. Ejaz Ahmed and Dr. Aqil Burney for their valuable comments and suggestions on the original manuscript.

Conflicts of Interest

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

References

[1] Wazwaz, A.-M. (2011) Linear and Nonlinear Integral Equations: Methods and Applications. Higher Education Press, Beijing. https://doi.org/10.1007/978-3-642-21449-3

[2] Elzaki, T.M. (2011) The New Integral Transform "Elzaki Transform". *Global Journal of Pure and Applied Mathematics*, **7**, 57-64.

[3] Abdelrahim Mahgob, M.M. and Elzaki, T.M. (2017) Elzaki Transform and Integro-Differential Equation with a Bulge Function. *IOSR Journal of Mathematics*, **11**, 2278-5728.

[4] Adam, B.A.A. (2015) A Comparative Study of Adomain Decomposition Method and the New Integral Transform "Elzaki Transform". *International Journal of Applied Mathematical Research*, **4**, 8-14. https://doi.org/10.14419/ijamr.v4i1.3799

[5] Elzaki, T.M. and Ezaki, S.M. (2011) On the Solution of Integro-Differential Equation Systems by Using ELzaki Transform. *Global Journal of Mathematical Science: Theory and Practical*, **3**, 13-23. http://www.irphouse.com

[6] Elzaki, T.M., Elzaki, S.M. and Elnour, E.A. (2012) On the New Integral Transform "ELzaki Transform" Fundamental Properties Investigations and Applications. *Global Journal of Mathematical Science: Theory and Practical*, **4**, 15-23.

[7] Jung, K. and Kim, H. (2014) The Practical Formulas for Differentiation of Integral Transforms. *International Journal of Mathematical Analysis*, **8**, 471-480. https://doi.org/10.12988/ijma.2014.4238

[8] Kim, H. (2014) A Note on the Shifting Theorems for the Elzaki Transform. *International Journal of Mathematical Analysis*, **8**, 481-488. https://doi.org/10.12988/ijma.2014.4248

[9] Ocvirk, E. (2007) An Application of Romberg Extrapolation on Quadrature Method for Solving Linear Volterra Integral Equations of the Second Kind. *Applied Mathematics and Computation*, **194**, 389-393. https://doi.org/10.1016/j.amc.2007.04.043

[10] Odibat, Z.M. (2008) Differential Transform Method for Solving Volterra Integral Equation with Separable Kernels. *Mathematical and Computer Modelling*, **48**, 1144-1149. https://doi.org/10.1016/j.mcm.2007.12.022

[11] Rashidinia, J. and Zarebnia, M. (2007) Solution of a Volterra Integral Equation by the Sinc-Collocation Method. *Journal of Computational and Applied Mathematics*, **206**, 801-813. https://doi.org/10.1016/j.cam.2006.08.036

[12] Song, Y. and Kim, H. (2014) The Solution of Volterra Integral Equation of the Second Kind by Using the Elzaki Transform. *Applied Mathematical Sciences*, **8**, 525-530. https://doi.org/10.12988/ams.2014.312715

[13] Elzaki, T.M. and Ezaki, S.M. (2011) Solution of Integro-Differential Equations by Using ELzaki Transform. *Global Journal of Mathematical Science*: *Theory and Practical*, **3**, 1-11. http://www.irphouse.com

[14] Makroglou, A. and Konstantinides, D. (2006) Numerical Solution of a System of Two First Order Volterra Integro-Differential Equations Arising in Ultimate Ruin Theory.

[15] Khezerloo, M. (2012) Existence and Uniqueness of Solution of Volterra Integral Equations. *International Journal of Industrial Mathematics*, **4**, 69-76.

[16] Jalalvand, M., Jazbi, B. and Mokhtarzadeh, M.R. (2013) A Finite Difference Method for the Smooth Solution of Linear Volterra Integral Equations. *The International Journal of Nonlinear Analysis and Applications* (*IJNAA*), **4**, 1-10.

[17] Rahman, M.M. (2013) Numerical Solutions of Volterra Integral Equations Using Galerkin Method with Hermite Polynomials. *Proceedings of the* 2013 *International Conference on Applied Mathematics and Computational Methods in Engineering*, Rhodes Island, 16-19 July 2013, 276-281.

[18] Wazwaz, A.-M. (2015) A First Course in Integral Equations. Second Edition, World Scientific Publishing Company, Singapore. https://doi.org/10.1016/j.jneuroim.2016.09.018

[19] Mirzaee, F. (2010) Numerical Computational Solution of the Linear Volterra Integral Equations System via Rationalized Haar Functions. *Journal of King Saud University—Science*, **22**, 265-268. https://doi.org/10.1016/j.jksus.2010.05.010

Journal of Applied Mathematics and Physics, 2019, 7, 1893-1910
http://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Two-Step Hybrid Block Method for Solving Nonlinear Jerk Equations

Bothayna S. H. Kashkari[1], Sadeem Alqarni[2,3]

[1]Department of Mathematics, Faculty of Science, University of Jeddah, Jeddah, KSA
[2]Department of Mathematics, Faculty of Science, King Abdulaziz University, Jeddah, KSA
[3]Department of Mathematics, Faculty of Science, Al-Baha University, Al-Baha, KSA
Email: bskashkari@uj.edu.sa, saalqarni@bu.edu.sa

How to cite this paper: Kashkari, B.S.H. and Alqarni, S. (2019) Two-Step Hybrid Block Method for Solving Nonlinear Jerk Equations. *Journal of Applied Mathematics and Physics*, **7**, 1893-1910.
https://doi.org/10.4236/jamp.2019.78130

Received: July 24, 2019
Accepted: August 23, 2019
Published: August 26, 2019

Abstract

In this paper, a block method with one hybrid point for solving Jerk equations is presented. The hybrid point is chosen to optimize the local truncation errors of the main formulas for the solution and the derivative at the end of the block. Analysis of the method is discussed, and some numerical examples show that the proposed method is efficient and accurate.

Keywords

Nonlinear Jerk Equations, Hybrid Block Method, Zero Stable, Consistent, Convergent

1. Introduction

Jerk is the rate of acceleration change in physics; that is, the time derivative of acceleration, and as such the second velocity derivative, or the third time position derivative. The jerk is important in several mechanics and acoustics applications. The Jerk vector is here resolved into tangential-normal and radial-transverse components for planar motion, and the normal component is expressed as an affine differential invariant recognized as the aberrancy. Several geometric properties of the Jerk vector are established for plane motion using known aberrancy properties of curves [1].

Nonlinear third-order differential equations, known as nonlinear Jerk equations, involving the third temporal displacement derivative, are of great interest in analyzing some structures which exhibit rotating and translating movements, such as robots or machine tools, where excessive Jerk leads to accelerated wear of transmissions and bearing elements, noisy operations and large contouring errors in discontinuities (such as corners) in the machining path [2].

Many authors have studied the numerical solutions of the Jerk equation, harmonic balance approach to periodic solutions is used in [3], in [4] they have written the high-order ordinary differential equation in terms of its differential invariants. New algorithm for the numerical solutions of nonlinear third-order differential equations was used jacobi-gauss collocation method in [5], He's variational iteration method was used in [6] for nonlinear Jerk equations. Modified harmonic balance method was used for nonlinear Jerk equations in [7]. In this paper, we consider a Jerk equation of the form

$$y''' = J\left(y, y', y''\right). \tag{1}$$

With initial conditions $y(0)=0, y'(0)=B, y''(0)=0$. The most general function of Jerk, which is invariant in time-reversal and space-reversal and only nonlinear cubic, can be written as

$$y''' + \alpha y y' y'' + \beta y' y''^2 + \delta y^2 y' + \epsilon y'^3 + \gamma y' = 0, \tag{2}$$

where the parameters are $\alpha, \beta, \delta, \epsilon$ and γ is constants. The current work is motivated by optimizing local truncation errors in order to find a hybrid point in a two-step block method to have the most accurate solution for Jerk Equation (1). We organize this paper as follows: The next section illustrates the method derivation, Section 3 presents the analysis of the method involving order four, Section 4 presents the numerical examples showing the productivity of the new technique when it is contrasted with the different strategies proposed in the scientific writing.

2. Methodology

To derive two steps hybrid block method with one off-step point we have a polynomial of degree 6 as follows:

$$y(x) \simeq p(x) = \sum_{j=0}^{6} a_j x^j, \tag{3}$$

with third derivative given by

$$y'''(x) \simeq p'''(x) = \sum_{j=3}^{6} j(j-1)(j-2) a_j x^{j-3}. \tag{4}$$

Substituting Equation (4) into Equation (3) gives

$$f\left(x, y, y', y''\right) = \sum_{j=3}^{6} j(j-1)(j-2) a_j x^{j-3}. \tag{5}$$

By interpolating Equation (3) at $x_{m+j}, j=0, r, 1$ and collocating Equation (5) at $x_{m+j}, j=0, r, 1, 2$ we get a system of equations written in the matrix form

$$\begin{bmatrix} 1 & x_m & x_m^2 & x_m^3 & x_m^4 & x_m^5 & x_m^6 \\ 1 & x_{m+r} & x_{m+r}^2 & x_{m+r}^3 & x_{m+r}^4 & x_{m+r}^5 & x_{m+r}^6 \\ 1 & x_{m+1} & x_{m+1}^2 & x_{m+1}^3 & x_{m+1}^4 & x_{m+1}^5 & x_{m+1}^6 \\ 0 & 0 & 0 & 6 & 24x_m & 60x_m^2 & 120x_m^3 \\ 0 & 0 & 0 & 6 & 24x_{m+r} & 60x_{m+r}^2 & 120x_{m+r}^3 \\ 0 & 0 & 0 & 6 & 24x_{m+1} & 60x_{m+1}^2 & 120x_{m+1}^3 \\ 0 & 0 & 0 & 6 & 24x_{m+2} & 60x_{m+2}^2 & 120x_{m+2}^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} = \begin{bmatrix} y_m \\ y_{m+r} \\ y_{m+1} \\ f_m \\ f_{m+r} \\ f_{m+1} \\ f_{m+2} \end{bmatrix},$$

Solving the above system (by using Gaussian elimination method for solving the system of linear equations [8]) gives us the coefficients of the polynomial a_j, $j=0,1,\cdots,6$.

By making the substitution $x=x_m+th$, the polynomial in Equation (3) may be written in the form:

$$p\left(x_m+th\right)=\alpha_0 y_m+\alpha_r y_{m+r}+\alpha_1 y_{m+1}+h^3\left(\beta_0 f_m+\beta_r f_{m+r}+\beta_1 f_{m+1}+\beta_2 f_{m+2}\right) \quad (6)$$

where

$$\alpha_0=-\frac{(r-t)(t-1)}{r},$$

$$\alpha_r=\frac{t(t-1)}{r(r-1)},$$

$$\alpha_1=\frac{t(r-t)}{r-1},$$

$$\beta_0=-\frac{h^3}{240r}\Big(t(t-1)\Big(r^4-8r^3+22r^2-2rt^3+13rt^2 -27rt-5r+t^4-5t^3+5t^2+5t\Big)\Big),$$

$$\beta_r=-\frac{h^3}{120r\left(r^2-3r+2\right)}\Big(t(t-1)\Big(r^4-5r^3+5r^2+5r-t^4+5t^3-5t^2-5t\Big)\Big),$$

$$\beta_1=-\frac{h^3}{120(r-1)}\Big(t(t-1)\Big(-r^4+5r^3+5r^2+2rt^3-8rt^2 -8rt-3r-t^4+3t^3+3t^2+3t\Big)\Big),$$

$$\beta_2=\frac{h^3}{240(r-2)}\Big(t(t-1)\Big(-r^4+2r^3+2r^2+2rt^3-3rt^2 -3rt-r-t^4+t^3+t^2+t\Big)\Big). \quad (7)$$

Now, by evaluating the solution approximation at the point x_{m+2}:

$$\begin{aligned} y_{m+2}=&-\frac{r-2}{r}y_m+\frac{2}{r(r-1)}y_{m+r}+\frac{2r-4}{r-1}y_{m+1}\\ &-\frac{h^3\left(r^4-8r^3+22r^2-23r+6\right)}{120r}f_m+\frac{h^3\left(-r^2+2r+3\right)}{60r}f_{m+r}\\ &-\frac{h^3\left(-r^3+4r^2+9r-26\right)}{60}f_{m+1}+\frac{h^3\left(-r^3+2r+1\right)}{120}f_{m+2}. \end{aligned} \quad (8)$$

Assess the approximation at the point of first derivative x_{m+2}:

$$\begin{aligned} hy'_{m+2}=&-\frac{r-3}{r}y_m+\frac{3}{r(r-1)}y_{m+r}+\frac{r-4}{r-1}y_{m+1}\\ &-\frac{h^3\left(3r^4-24r^3+66r^2-67r+12\right)}{240r}f_m\\ &-\frac{h^3\left(r^4-5r^3+5r^2+5r-4\right)}{(40r(r^2-3r+2))}f_{m+r} \end{aligned}$$

$$
\begin{aligned}
&-\frac{h^3\left(-3r^4+15r^3+15r^2-137r+116\right)}{120(r-1)}f_{m+1}\\
&+\frac{h^3\left(-r^4+2r^3+2r^2+3r-12\right)}{80(r-2)}f_{m+2}.
\end{aligned}
\tag{9}
$$

Evaluate the second derivative approximation at the point x_{m+2}:

$$
\begin{aligned}
h^2y''_{m+2}=&\frac{2}{r}y_m+\frac{2}{r(r-1)}y_{m+r}-\frac{2}{r-1}y_{m+1}\\
&+\frac{h^3\left(-r^4+8r^3-22r^2+18r+5\right)}{120r}f_m\\
&-\frac{h^3\left(r^4-5r^3+5r^2+5r+5\right)}{(60r(r^2-3r+2))}f_{m+r}\\
&-\frac{h^3\left(-r^4+5r^3+5r^2-75r+77\right)}{60(r-1)}f_{m+1}\\
&+\frac{h^3\left(-r^4+2r^3+2r^2+42r-81\right)}{120(r-2)}f_{m+2}.
\end{aligned}
\tag{10}
$$

We choose to optimize the local truncation errors in the Equation (8), Equation (9) and Equation (10) to determine appropriate values for r Equation (8), Equation (9) and Equation (10). This choice at the end of the block y_{m+2}, y'_{m+2} and y''_{m+2}, which result respectively in

$$
\begin{aligned}
&\mathcal{L}\left(y\left(x_{m+2};h\right)\right)\\
&=-\frac{h^7y^{(7)}\left(x_m\right)}{10080}\left(3r^5-18r^4+24r^3+24r^2-53r+10\right)\\
&\quad-\frac{h^8y^{(8)}\left(x_m\right)}{100800}\left(9r^6-33r^5-47r^4+205r^3+51r^2-327r+42\right),\\
&\mathcal{L}\left(y'\left(x_{m+2};h\right)\right)\\
&=\frac{h^7y^{(7)}\left(x_m\right)}{20160}\left(-9r^5+54r^4-72r^3-72r^2+117r+44\right)\\
&\quad+\frac{h^8y^{(8)}\left(x_m\right)}{201600}\left(-27r^6+99r^5+141r^4-615r^3-237r^2+729r+556\right),\\
&\mathcal{L}\left(y''\left(x_{m+2};h\right)\right)\\
&=-\frac{h^7y^{(7)}\left(x_m\right)}{3360}\left(r^5-6r^4+8r^3+8r^2+8r-41\right)\\
&\quad-\frac{h^8y^{(8)}\left(x_m\right)}{100800}\left(9r^6-33r^5-47r^4+205r^3+205r^2+135r-1195\right).
\end{aligned}
\tag{11}
$$

Determine the r values equating to zero the h^7 coefficients in the local truncation error formulas in Equation (11), and we obtain the system

$$
\begin{aligned}
&-3r^5+18r^4-24r^3-24r^2+53r-10=0,\\
&-9r^5+54r^4-72r^3-72r^2+117r+44=0,
\end{aligned}
$$

$$-r^5+6r^4-8r^3-8r^2-8r+41=0. \tag{12}$$

By solving Equation (12) and substituting r's in Equation (11) we can choose the value r that gives the least truncation errors

$$r=\frac{161757+\sqrt{36610681841}}{199636}.$$

Substituting r in the local truncation errors gives:

$$\mathcal{L}\left(y\left(x_m;h\right)\right)\approx\frac{2h^7y^{(7)}}{211915}+O\left(h^8\right)$$

$$\mathcal{L}\left(y'\left(x_m;h\right)\right)\approx-\frac{17h^8y^{(8)}}{107192}+O\left(h^9\right)$$

$$\mathcal{L}\left(y''\left(x_m;h\right)\right)\approx\frac{-26h^7y^{(7)}}{84685}+O\left(h^8\right)$$

To obtain a two-step hybrid block solution method for solving Equation (1), we evaluate $p'(x), p''(x)$ at the points x_m, x_{m+r}, x_{m+1}, we get the following block of six equations

$$\begin{aligned}hy'_m=&-\frac{r+1}{r}y_m-\frac{1}{r(r-1)}y_{m+r}+\frac{r}{r-1}y_{m+1}\\&+\frac{h^3\left(r^3-8r^2+22r-5\right)}{240}f_m\\&+\frac{h^3\left(r^3-5r^2+5r+5\right)}{120\left(r^2-3r+2\right)}f_{m+r}\\&+\frac{h^3r\left(-r^3+5r^2+5r-3\right)}{120(r-1)}f_{m+1}\\&-\frac{h^3r\left(-r^3+2r^2+2r-1\right)}{240(r-2)}f_{m+2},\end{aligned}$$

$$\begin{aligned}h^2y''_m=&\frac{2}{r}y_m+\frac{2}{r(r-1)}y_{m+r}-\frac{2}{r-1}y_{m+1}\\&-\frac{h^3\left(r^4-8r^3+22r^2-28r+10\right)}{120r}f_m\\&-\frac{h^3\left(r^4-5r^3+5r^2+5r-10\right)}{60r\left(r^2-3r+2\right)}f_{m+r}\\&-\frac{h^3\left(-r^4+5r^3+5r^2-35r+22\right)}{60(r-1)}f_{m+1}\\&+\frac{h^3\left(-r^4+2r^3+2r^2-8r+4\right)}{120(r-2)}f_{m+2},\end{aligned}$$

$$\begin{aligned}hy'_{m+r}=&\frac{r-1}{r}y_m+\frac{2r-1}{r(r-1)}y_{m+r}-\frac{r}{r-1}y_{m+1}\\&+\frac{h^3\left(2r^4-13r^3+28r^2-22r+5\right)}{240}f_m\end{aligned}$$

$$+\frac{h^3\left(4r^3-15r^2+10r+5\right)}{120(r-2)}f_{m+r}$$
$$+\frac{h^3r\left(-2r^3+7r^2+2r-3\right)}{120}f_{m+1}$$
$$+\frac{h^3r\left(2r^4-5r^3+2r^2+2r-1\right)}{240(r-2)}f_{m+2},$$

$$h^2y''_{m+r}=\frac{2}{r}y_m+\frac{2}{r(r-1)}y_{m+r}-\frac{2}{r-1}y_{m+1}$$
$$+\frac{h^3\left(4r^4-22r^3+38r^2-22r+5\right)}{120r}f_m$$
$$-\frac{h^3\left(-14r^4+55r^3-55r^2+5r+5\right)}{60r\left(r^2-3r+2\right)}f_{m+r}$$
$$-\frac{h^3\left(4r^4-15r^3+5r^2+5r-3\right)}{60(r-1)}f_{m+1}$$
$$+\frac{h^3\left(4r^4-8r^3+2r^2+2r-1\right)}{120(r-2)}f_{m+2},$$

$$hy'_{m+1}=-\frac{r-1}{r}y_m+\frac{1}{r(r-1)}y_{m+r}+\frac{r-2}{r-1}y_{m+1}$$
$$-\frac{h^3\left(r^4-8r^3+22r^2-21r+6\right)}{240r}f_m$$
$$+\frac{h^3\left(-r^2+2r+3\right)}{120r}f_{m+r}$$
$$-\frac{h^3\left(-r^3+4r^2+9r-8\right)}{120}f_{m+1}$$
$$-\frac{h^3\left(r^3-2r+1\right)}{240}f_{m+2},$$

$$h^2y''_{m+1}=\frac{2}{r}y_m+\frac{2}{r(r-1)}y_{m+r}-\frac{2}{r-1}y_{m+1}$$
$$-\frac{h^3\left(r^4-8r^3+22r^2-28r+10\right)}{120r}f_m$$
$$-\frac{h^3\left(r^4-5r^3+5r^2+5r-10\right)}{60r(r^2-3r+2)}f_{m+r}$$
$$-\frac{h^3\left(-r^4+5r^3+5r^2-35r+22\right)}{60(r-1)}f_{m+1}$$
$$+\frac{h^3\left(-r^4+2r^3+2r^2-8r+4\right)}{120(r-2)}f_{m+2}.$$ (13)

3. Characteristics of the Method

This section is presented the basic properties of the main method and analyzed it

to establish their validity.

3.1. Order of the Method

We can rewrite the hybrid block method in the form

$$AY_m = hBY'_m + h^2CY''_m + h^3VF_m \tag{14}$$

where A, B, C, V are matrices of coefficients of dimensions 9×4,

$$Y_m = \left(y_m, y_{m+r}, y_{m+1}, y_{m+2}\right)^{\mathrm{T}},$$

$$Y'_m = \left(y'_m, y'_{m+r}, y'_{m+1}, y'_{m+2}\right)^{\mathrm{T}},$$

$$Y''_m = \left(y''_m, y''_{m+r}, y''_{m+1}, y''_{m+2}\right)^{\mathrm{T}},$$

$$F_m = \left(f_m, f_{m+r}, f_{m+1}, f_{m+2}\right)^{\mathrm{T}}.$$

If $z(x)$ is a sufficiently differentiable function, the linear difference operator $\mathcal{L}$ associated with the implicit two-step block hybrid method is considered in Equation (8), Equation (9), Equation (10), Equation (13), that is given

$$\begin{aligned}\mathcal{L}\left[z(x_m);h\right] = &\sum_j \kappa_j z(x_m + jh) - h\tau_j z'(x_m + jh)\\ &- h^2\gamma_j z''(x_m + jh) - h^3\xi_j z'''(x_m + jh)\end{aligned} \tag{15}$$

$j = 0, r, 1, 2$ where the $\kappa_j, \tau_j, \gamma_j$ and ξ_j are respectively the vector columns of the matrices A, B, C and V.

Expanding $z(x_m + jh), z'(x_m + jh), z''(x_m + jh)$ and $z'''(x_m + jh)$ in Taylor series about x_m we get

$$\mathcal{L}\left[z(x_m);h\right] = C_0 y(x_m) + C_1 h y'(x_m) + C_2 h^2 y''(x_m) + \cdots + C_q h^q y^{(q)}(x_m) + \cdots \tag{16}$$

with $C_0 = C_1 = C_2 = \cdots = C_6 = 0$ and

$$\begin{aligned}C_7 = \big(&9.437739543552612\times 10^{-6}, -9.617040306545119\times 10^{-20},\\ &-3.070200924962517\times 10^{-4}, -4.913128762599697\times 10^{-4},\\ &2.144715335032775\times 10^{-3}, 8.823816935161847\times 10^{-5},\\ &-4.504524331158401\times 10^{-4}, 3.184040914146728\times 10^{-3},\\ &-1.141813120360736\times 10^{-2}\big)^{\mathrm{T}}\end{aligned}$$

Note that the proposed method has order $p = 4$ at least [9].

3.2. Zero Stability

We can write the method as a vector As $h \to 0$ in Equation (14), we can write the method as a vector form.

$A_0 Y_m - A_1 Y_{m-1} = 0$ where

$$Y_m = \left(y_{m+2}, y_{m+1}, y_{m+r}\right)^{\mathrm{T}}, Y_{m-1} = \left(y_m, y_{m-1}, y_{m+r-1}\right)^{\mathrm{T}}$$

$$A_0 = \begin{bmatrix} 1 & 0.6017954608298602 & -1.471021404061736 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 0.1307740567681242 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

The first characteristic polynomial

$\rho(z) = \det[A_0 z - A_1] = (z-1)z^2 = 0$, the roots of polynomials are $z_1 = z_2 = 0; z_3 = 1$. Hence the block method is zero -stable [9].

3.3. Consistency

The block method has order $p = 4$, in case $p \geq 1$, this is a sufficient condition to be consistent with the associated block method [10].

3.4. Convergence

We can establish the convergence of the two-step with three points hybrid block method if and only if it is consistent and zero stable [11].

3.5. Region of Absolute Stability

As we mentioned earlier, zero-stability is a concept of the numerical method behavior for $h \to 0$. To decide whether a numerical method will produce good results with a given value of $h > 0$, we need a concept of stability that is different from zero-stability. In most numerical methods intended to solve problems of third order, the stability properties are usually analyzed by considering the linear equation given by the Dalquist test [11].

$$y''' = -\lambda^3 y(x) \tag{17}$$

This problem has bounded solutions for $\lambda \geq 0$ that tend to zero for $x \to \infty$. We will define the region where the numerical method reproduces the manner of the exact solutions. Let us explain the procedure for obtaining such a region. Our method has nine equations in which there are four different terms of first derivatives: $y'_m, y'_{m+r}, y'_{m+1}, y'_{m+2}$, and second derivative: $y''_m, y''_{m+r}, y''_{m+1}, y''_{m+2}$ and one intermediate values y_{m+r}. Let us depict the procedure to gain such a region [12], We eliminated these terms from the equations system by using mathematica, and get a recurrence equation in the terms y_m, y_{m+1}, y_{m+2}. This recurrence equation reads

$$P(z)y_{m+2} + Q(z)y_{m+1} + S(z)y_m = 0, \tag{18}$$

where $z = \lambda h$, and

$$\begin{aligned} P(z) &= -1.433653893592262 \times 10^{109} + 5.520254762660736 \times 10^{107} z^3 \\ &\quad + 3.674797669760869 \times 10^{106} z^6 \\ Q(z) &= 5.734615574369057 \times 10^{109} - 8.898486741828162 \times 10^{108} z^3 \\ &\quad + 3.679066781172749 \times 10^{107} z^6 \\ S(z) &= -4.300961680776793 \times 10^{109} + 2.867307787184524 \times 10^{109} z \\ &\quad - 1.211231358386333 \times 10^{108} z^3 - 6.261663213347228 \times 10^{107} z^4 \\ &\quad + 4.300976789540145 \times 10^{106} z^6 \end{aligned} \tag{19}$$

Through its characteristic equation, we study the extent boundedness of their solutions to define the region of stability. The root causes of the characteristic equation

$$\begin{aligned}&\left(-1.433653893592262\times10^{109}+5.520254762660736\times10^{107}z^3\right.\\&\left.+3.674797669760869\times10^{106}z^6\right)x^2+\left(5.734615574369057\times10^{109}\right.\\&\left.-8.898486741828162\times10^{108}z^3+3.679066781172749\times10^{107}z^6\right)x\\&+\left(-4.300961680776793\times10^{109}+2.867307787184524\times10^{109}z\right.\\&\left.-1.211231358386333\times10^{108}z^3-6.261663213347228\times10^{107}z^4\right.\\&\left.+4.300976789540145\times10^{106}z^6\right)=0\end{aligned}\tag{20}$$

are

$$\begin{aligned}x_1=&-\frac{2}{4509981306728599z^6+67748616455391176z^3-1759485240133103872}\\&\times(\sqrt{(}121468655749357324985383025377280z^{12}+866454506980595547423584571883 52z^{10}\\&-6085569311397395324399777440858112z^9-2666041011195242645961971549077504z^7\\&+1250542873217974968358810166031810 56z^6-934044547739805699736080593704714 24z^4\\&-9367413665827868061177950035330990 08z^3+1547894155123122963374310449590829056z\\&+7739470775615676797931639912047575 04)\\&-546043788941838592z^3+22576103367311352z^6+3518970480266213376)\end{aligned}$$

$$\begin{aligned}x_2=&\frac{2}{4509981306728599z^6+67748616455391176z^3-1759485240133103872}\\&\times(\sqrt{(}121468655749357324985383025377280z^{12}+866454506980595547423584571883 52z^{10}\\&-6085569311397395324399777440858112z^9-2666041011195242645961971549077504z^7\\&+1250542873217974968358810166031810 56z^6-934044547739805699736080593704714 24z^4\\&-9367413665827868061177950035330990 08z^3+1547894155123122963374310449590829056z\\&+7739470775615676797931639912047575 04)\\&+546043788941838592z^3-22576103367311352z^6-3518970480266213376)\end{aligned}$$

The roots of the characteristic equation must be less than 1, for the method to be stable. The stability region for the method has shown in **Figure 1**.

4. Numerical Examples

Example 1: Consider the Jerk equation in the following form:

$$y'''=-y'+yy'y''\tag{21}$$

With initial conditions: $y(0)=0, y'(0)=B, y''(0)=0$, and exact solution:

$$y(x)=\frac{B}{\Omega}\sin(\Omega x)+\frac{B}{96\Omega^3}\left(\left(-9B^2-48+48\Omega^2\right)\sin(\Omega x)-B^2\sin(3\Omega x)\right)$$

where $\Omega=\frac{1}{2}\sqrt{B^2+4}$.

Tables 1-3 show the absolute Errors at $h=0.125, 0.0125$ and $B=0.2, 0.3, 0.4$ respectively. Example 1 was solved in [13], the comparison be-

tween our method and result in [13] is proposed in Table 4 and Table 5. It is seen in Figures 2-4 the method gives a good approximation in the interval $x \in [0,50]$ and $B = 0.2, 0.3, 0.4$ respectively.

Example 2: Consider the Jerk equation in the following form:

$$y''' = -y' - y'y''^2 \tag{22}$$

With initial conditions: $y(0) = 0, y'(0) = B, y''(0) = 0$, and exact solution:

$$y(x) = \frac{B}{\Omega}\sin(\Omega x) + \frac{B}{96\Omega^3}\left(\left(-9B^2\Omega^2 - 48 + 48\Omega^2\right)\sin(\Omega x) + \left(12\Omega^3 B^2 + 48\Omega - 48\Omega^3\right)x\cos(\Omega x) - B^2\Omega^2\sin(3\Omega x)\right)$$

where $\Omega = 2\sqrt{\frac{1}{4 - B^2}}$.

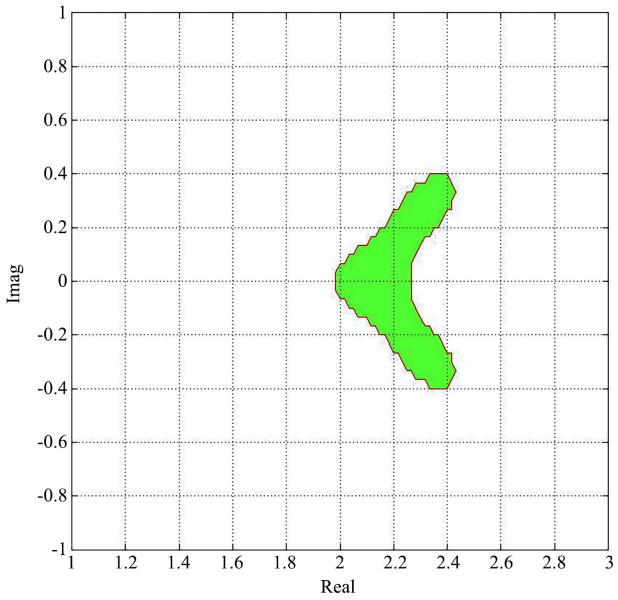


Figure 1. Region of absolute stability.

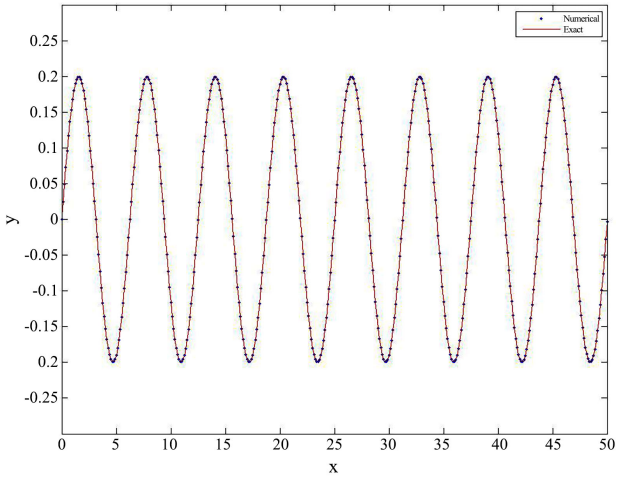


Figure 2. The solution of Example 1 at $h = 0.125$, $B = 0.2$.

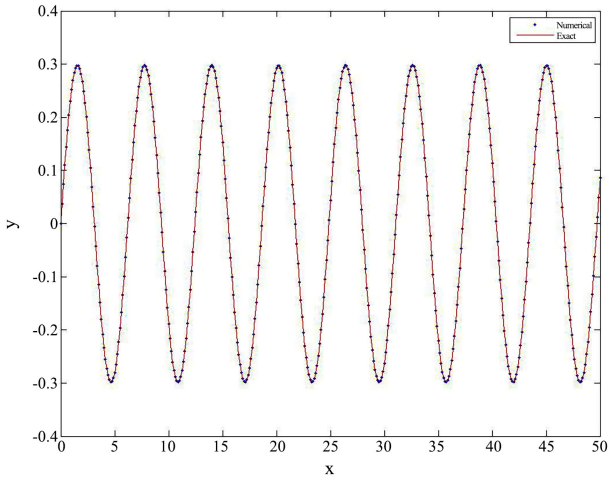


Figure 3. The solution of Example 1 at $h = 0.125$, $B = 0.3$.

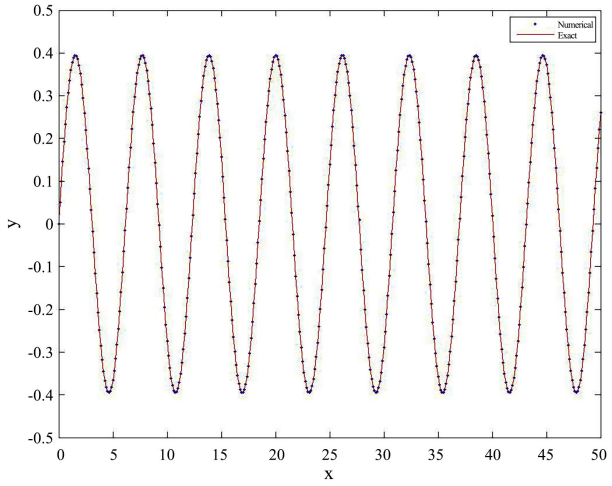


Figure 4. The solution of Example 1 at $h = 0.125$, $B = 0.4$.

Table 1. Absolute error of numerical solutions at $B = 0.2$.

X	Numerical Solution in Our method at $h = 0.125$	Exact Solution at $h = 0.125$	Error	Numerical Solution in Our method at $h = 0.0125$	Exact Solution at $h = 0.0125$	Error
0	0	0	0	0	0	0
0.125	0.024934	0.024934	4.6952×10^{-13}	0.024934	0.0249349	4.5195×10^{-11}
0.25	0.049480	0.049480	1.09357×10^{-9}	0.049480	0.049480	1.39199×10^{-9}
0.375	0.073253	0.073253	9.03272×10^{-9}	0.073253	0.073253	9.87559×10^{-9}
0.5	0.095881	0.095881	3.59222×10^{-8}	0.095881	0.095881	3.81834×10^{-8}
0.625	0.117008	0.117008	9.9039×10^{-8}	0.117008	0.117008	1.037×10^{-7}
0.75	0.136300	0.1363	2.14523×10^{-7}	0.136300	0.1363	2.23488×10^{-7}
0.875	0.153453	0.153452	3.91443×10^{-7}	0.153453	0.153452	4.0688×10^{-7}
1	0.168191	0.168191	6.24133×10^{-7}	0.168191	0.168191	6.49333×10^{-7}
1.125	0.180281	0.18028	8.91573×10^{-7}	0.180281	0.18028	9.29984×10^{-7}
1.25	0.189526	0.189525	1.15844×10^{-6}	0.189526	0.189525	1.21443×10^{-6}
1.375	0.195778	0.195777	1.38459×10^{-6}	0.195779	0.195777	1.46242×10^{-6}

Continued

1.5	0.198937	0.198935	1.53404×10^{-6}	0.198937	0.198935	1.63819×10^{-6}
1.625	0.198949	0.198948	1.58573×10^{-6}	0.198949	0.198948	1.72019×10^{-6}
1.75	0.195816	0.195815	1.53899×10^{-6}	0.195817	0.195815	1.70703×10^{-6}
1.875	0.189589	0.189587	1.41336×10^{-6}	0.189589	0.189587	1.61747×10^{-6}
2	0.180367	0.180365	1.24351×10^{-6}	0.180367	0.180365	1.48465×10^{-6}

Table 2. Absolute error of numerical solutions at $B=0.3$.

X	Numerical Solution in Our method at $h=0.125$	Exact Solution at $h=0.125$	Error	Numerical Solution in Our method at $h=0.0125$	Exact Solution at $h=0.0125$	Error
0	0	0	0	0	0	0
0.125	0.037402	0.037402	6.28175×10^{-10}	0.037402	0.037402	3.43214×10^{-10}
0.25	0.074220	0.074220	1.24193×10^{-8}	0.074220	0.074220	1.05686×10^{-8}
0.375	0.109879	0.109878	7.99073×10^{-8}	0.109878	0.109878	7.55001×10^{-8}
0.5	0.143814	0.143814	2.98246×10^{-7}	0.143814	0.143814	2.89694×10^{-7}
0.625	0.17549	0.17549	7.99279×10^{-7}	0.17549	0.17549	7.86251×10^{-7}
0.75	0.204399	0.204397	1.70916×10^{-6}	0.204399	0.204397	1.69292×10^{-6}
0.875	0.230073	0.23007	3.09552×10^{-6}	0.230073	0.23007	3.07827×10^{-6}
1	0.252093	0.252088	4.91735×10^{-6}	0.252093	0.252088	4.90448×10^{-6}
1.125	0.270095	0.270088	7.01179×10^{-6}	0.270095	0.270088	7.00958×10^{-6}
1.25	0.283778	0.283769	9.11212×10^{-6}	0.283778	0.283769	9.13002×10^{-6}
1.375	0.292913	0.292902	0.000010	0.292913	0.292902	0.000010
1.5	0.297344	0.297332	0.000012	0.297344	0.297332	0.000012
1.625	0.296996	0.296984	0.000012	0.296996	0.296984	0.000012
1.75	0.291876	0.291864	0.000012	0.291876	0.291864	0.000012
1.875	0.28207	0.282059	0.000011	0.282071	0.282059	0.000011
2	0.267745	0.267734	0.000010	0.267745	0.267734	0.000010

Table 3. Absolute error of numerical solutions at $B=0.4$.

X	Numerical Solution in Our method at $h=0.125$	Exact Solution at $h=0.125$	Error	Numerical Solution in Our method at $h=0.0125$	Exact Solution at $h=0.0125$	Error
0	0	0	0	0	0	0
0.125	0.049869	0.049869	2.46607×10^{-9}	0.049869	0.049869	1.4462×10^{-9}
0.25	0.098960	0.098960	5.11558×10^{-8}	0.098961	0.098961	4.45237×10^{-8}
0.375	0.146501	0.146501	3.34007×10^{-7}	0.146501	0.146501	3.17797×10^{-7}
0.5	0.191739	0.191738	1.25262×10^{-6}	0.191739	0.191738	1.21922×10^{-6}
0.625	0.233947	0.233943	3.36066×10^{-6}	0.233947	0.233943	3.30604×10^{-6}
0.75	0.272436	0.272428	7.18675×10^{-6}	0.272436	0.272428	7.10926×10^{-6}
0.875	0.306566	0.306553	0.000013	0.306566	0.306553	0.000012
1	0.33576	0.335739	0.000021	0.33576	0.335739	0.000020
1.125	0.359513	0.359483	0.000029	0.359513	0.359483	0.000029
1.25	0.377408	0.37737	0.000038	0.377408	0.37737	0.000037
1.375	0.389127	0.389082	0.000045	0.389127	0.389082	0.000045
1.5	0.39446	0.394409	0.000050	0.39446	0.394409	0.000050

Continued

1.625	0.393308	0.393256	0.000052	0.393308	0.393256	0.000052
1.75	0.385695	0.385643	0.000051	0.385695	0.385643	0.000051
1.875	0.371756	0.371707	0.000048	0.371756	0.371707	0.000048
2	0.351742	0.351698	0.000044	0.351742	0.351698	0.000044

Table 4. Comparison between the numerical solution in our method and the method in [13], $B = 0.2, h = 0.125$.

X	Numerical Solution in our method	Numerical Solution in [13]
0	0	0
0.125	0.024934	0.024934
0.25	0.049480	0.049480
0.375	0.073253	0.073254
0.5	0.095881	0.095885
0.625	0.117008	0.117019
0.75	0.1363	0.136327
0.875	0.153453	0.153508
1	0.168191	0.168294
1.125	0.180281	0.180453
1.25	0.189526	0.189796
1.375	0.195778	0.196178
1.5	0.198937	0.199498
1.625	0.198949	0.199706
1.75	0.195816	0.196797
1.875	0.189589	0.190817
2	0.180367	0.181859

Table 5. Comparsion between the numerical solution in our method and the method in [13], $B = 0.4, h = 0.125$.

X	Numerical Solution in our method	Numerical Solution in [13]
0	0	0
0.125	0.049869	0.049869
0.25	0.098960	0.098961
0.375	0.146501	0.146509
0.5	0.191739	0.191770
0.625	0.233947	0.234038
0.75	0.272436	0.272655
0.875	0.306566	0.307017
1	0.33576	0.336588
1.125	0.359513	0.360907
1.25	0.377408	0.379593
1.375	0.389127	0.392357
1.5	0.39446	0.398997
1.625	0.393308	0.399412
1.75	0.385695	0.393594
1.875	0.371756	0.381634
2	0.351742	0.363718

Tables 6-8 show the absolute Errors at $h = 0.125, 0.0125$ and $B = 0.2, 0.3, 0.4$ respectively. Example 2 was solved in [13], the comparison between our method and result in [13] is proposed in Table 9 and Table 10. It is seen in Figures 5-7 the method gives a good approximation in the interval $x \in [0, 50]$ and $B = 0.2, 0.3, 0.4$ respectively.

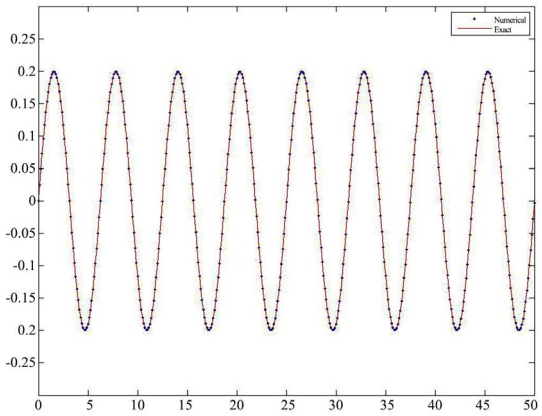


Figure 5. The solution of Example 2 at $h = 0.125$, $B = 0.2$.

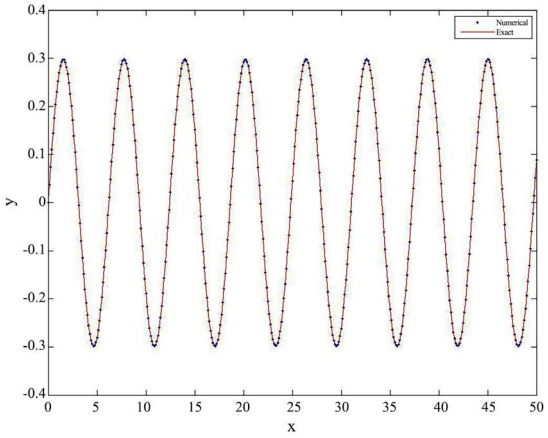


Figure 6. The solution of Example 2 at $h = 0.125$, $B = 0.3$.

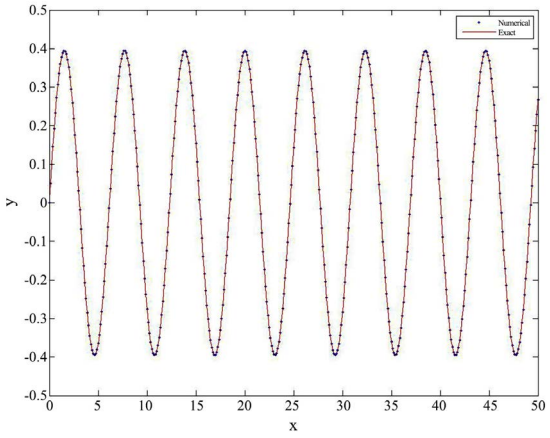


Figure 7. The solution of Example 2 at $h = 0.125$, $B = 0.4$.

Table 6. Absolute error of numerical solutions at $B = 0.2$.

X	Numerical Solution in Our method at $h = 0.125$	Exact Solution at $h = 0.125$	Error	Numerical Solution in Our method at $h = 0.0125$	Exact Solution at $h = 0.0125$	Error
0	0	0	0	0	0	0
0.125	0.024934	0.024934	3.89579×10^{-11}	0.024934	0.024934	8.68102×10^{-11}
0.25	0.049480	0.049480	2.37202×10^{-9}	0.049480	0.049480	2.68466×10^{-9}
0.375	0.073253	0.073253	1.83415×10^{-8}	0.073253	0.073253	1.92163×10^{-8}
0.5	0.095881	0.095881	7.26168×10^{-8}	0.095881	0.095881	7.49309×10^{-8}
0.625	0.117008	0.117008	2.01662×10^{-7}	0.117008	0.117008	2.06392×10^{-7}
0.75	0.1363	0.1363	4.44143×10^{-7}	0.1363	0.1363	4.53178×10^{-7}
0.875	0.153452	0.153452	8.29786×10^{-7}	0.153452	0.153452	8.45282×10^{-7}
1	0.168191	0.16819	1.3665×10^{-6}	0.168191	0.16819	1.39175×10^{-6}
1.125	0.18028	0.180278	2.03637×10^{-6}	0.18028	0.180278	2.07483×10^{-6}
1.25	0.189525	0.189522	2.79641×10^{-6}	0.189525	0.189522	2.85251×10^{-6}
1.375	0.195777	0.195773	3.59097×10^{-6}	0.195777	0.195773	3.66902×10^{-6}
1.5	0.198934	0.198929	4.36561×10^{-6}	0.198934	0.198929	4.47017×10^{-6}
1.625	0.198945	0.19894	5.08301×10^{-6}	0.198945	0.19894	5.21817×10^{-6}
1.75	0.195811	0.195805	5.73169×10^{-6}	0.195811	0.195805	5.90079×10^{-6}
1.875	0.189581	0.189575	6.32606×10^{-6}	0.189581	0.189575	6.53163×10^{-6}
2	0.180357	0.18035	6.89847×10^{-6}	0.180357	0.18035	7.14149×10^{-6}

Table 7. Absolute error of numerical solutions at $B = 0.3$.

X	Numerical Solution in Our method at $h = 0.125$	Exact Solution at $h = 0.125$	Error	Numerical Solution in Our method at $h = 0.0125$	Exact Solution at $h = 0.0125$	Error
0	0	0	0	0	0	0
0.125	0.037402	0.037402	9.40558×10^{-10}	0.037402	0.037402	6.72264×10^{-10}
0.25	0.074220	0.074220	2.25371×10^{-8}	0.074220	0.074220	2.07935×10^{-8}
0.375	0.109879	0.109878	1.53599×10^{-7}	0.109878	0.109878	1.49427×10^{-7}
0.5	0.143814	0.143814	5.88902×10^{-7}	0.143814	0.143814	5.80723×10^{-7}
0.625	0.17549	0.175489	1.61265×10^{-6}	0.17549	0.175489	1.60006×10^{-6}
0.75	0.204399	0.204395	3.52997×10^{-6}	0.204399	0.204395	3.51404×10^{-6}
0.875	0.230072	0.230066	6.57212×10^{-6}	0.230072	0.230066	6.55487×10^{-6}
1	0.252091	0.25208	0.000010	0.252091	0.25208	0.000010
1.125	0.27009	0.270074	0.000016	0.27009	0.270074	0.000016
1.25	0.283769	0.283747	0.000022	0.283769	0.283747	0.000022
1.375	0.292898	0.29287	0.000028	0.292898	0.29287	0.000028
1.5	0.297321	0.297287	0.000034	0.297321	0.297287	0.000034
1.625	0.296964	0.296924	0.000040	0.296964	0.296924	0.000040
1.75	0.291832	0.291786	0.000045	0.291832	0.291786	0.000045
1.875	0.282012	0.281962	0.000050	0.282013	0.281962	0.000050
2	0.267672	0.267617	0.000054	0.267672	0.267617	0.000055

Table 8. Absolute error of numerical solutions at $B = 0.4$.

X	Numerical Solution in Our method at $h = 0.125$	Exact Solution at $h = 0.125$	Error	Numerical Solution in Our method at $h = 0.0125$	Exact Solution at $h = 0.0125$	Error
0	0	0	0	0	0	0
0.125	0.049869	0.049869	3.86139×10^{-9}	0.049869	0.049869	2.91296×10^{-9}
0.25	0.098960	0.098960	9.62982×10^{-8}	0.098960	0.09896	9.01239×10^{-8}
0.375	0.146501	0.146501	6.62938×10^{-7}	0.146501	0.146501	6.47732×10^{-7}
0.5	0.191739	0.191736	2.55096×10^{-6}	0.191739	0.191736	2.51913×10^{-6}
0.625	0.233946	0.233939	6.99674×10^{-6}	0.233946	0.233939	6.9439×10^{-6}
0.75	0.272434	0.272419	0.000015	0.272434	0.272419	0.000015
0.875	0.306562	0.306533	0.000028	0.306561	0.306533	0.000028
1	0.335749	0.335702	0.000046	0.335749	0.335702	0.000046
1.125	0.359492	0.359422	0.000069	0.359492	0.359422	0.000069
1.25	0.37737	0.377275	0.000095	0.37737	0.377275	0.000095
1.375	0.389065	0.388942	0.000122	0.389065	0.388942	0.000122
1.5	0.394363	0.394214	0.000149	0.394363	0.394214	0.000149
1.625	0.393169	0.392994	0.000174	0.393169	0.392994	0.000174
1.75	0.385503	0.385307	0.000196	0.385504	0.385307	0.000196
1.875	0.371507	0.371289	0.000217	0.371507	0.371289	0.000217
2	0.351432	0.351193	0.000238	0.351432	0.351193	0.000238

Table 9. Comparsion between the numerical solution in our method and the method in [13], $B = 0.2, h = 0.125$.

X	Numerical Solution in our method	Numerical Solution in [13]
0	0	0
0.125	0.024934	0.024934
0.25	0.049480	0.049480
0.375	0.073253	0.073254
0.5	0.095881	0.095885
0.625	0.117008	0.117019
0.75	0.1363	0.136327
0.875	0.153452	0.153508
1	0.168191	0.168294
1.125	0.18028	0.180453
1.25	0.189525	0.189796
1.375	0.195777	0.196178
1.5	0.198934	0.199498
1.625	0.198945	0.199706
1.75	0.195811	0.196797
1.875	0.189581	0.190817
2	0.180357	0.181859

Table 10. Comparison between the numerical solution in our method and the method in [13], $B = 0.4$.

X	Numerical Solution in our method	Numerical Solution in [13]
0	0	0
0.125	0.049869	0.049869
0.25	0.098960	0.098961
0.375	0.146501	0.146509
0.5	0.191739	0.191770
0.625	0.233946	0.234038
0.75	0.272434	0.272655
0.875	0.306562	0.307017
1	0.335749	0.336588
1.125	0.359492	0.360907
1.25	0.37737	0.379593
1.375	0.389065	0.392357
1.5	0.394363	0.398997
1.625	0.393169	0.399412
1.75	0.385503	0.393594
1.875	0.371507	0.381634
2	0.351432	0.363718

5. Conclusion

A two-step with one hybrid point was proposed and proceeded as a self-starting method for solving nonlinear Jerk equations. We considered one hybrid point and specified for approximation after optimizing local truncation errors related to the main formula. Therefore, our method's good convergent and stability properties make it attractive for the numerical solution of nonlinear problems. The method presented is zero stable, consistent and convergence of four-algebraic order. The numerical results and figures show their efficiency and precision compared to other methods in the literature.

Conflicts of Interest

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

References

[1] Schot, S.H. (1978) Jerk: The Time Rate of Change of Acceleration. *American Journal of Physics*, **46**, 1090-1094. https://doi.org/10.1119/1.11504

[2] Anu, N. and Marinca, V. (2016) Approximate Analytical Solutions to Jerk Equations. *Dynamical Systems: Theoretical and Experimental Analysis*, Łódź, 7-10 December 2015, 169-176. https://doi.org/10.1007/978-3-319-42408-8_14

[3] Gottlieb, H. (2004) Harmonic Balance Approach to Periodic Solutions of Non-Linear

Jerk Equations. *Journal of Sound and Vibration*, **271**, 671-683. https://doi.org/10.1016/S0022-460X(03)00299-2

[4] Momoniat, E. and Mahomed, F. (2010) Symmetry Reduction and Numerical Solution of a Third-Order ODE from Thin Film Flow. *Mathematical and Computational Applications*, **15**, 709-719. https://doi.org/10.3390/mca15040709

[5] Bhrawy, A. and Abd-Elhameed, W. (2011) New Algorithm for the Numerical Solutions of Nonlinear Third-Order Differential Equations Using Jacobi-Gauss Collocation Method. *Mathematical Problems in Engineering*, **2011**, Article ID: 837218. https://doi.org/10.1155/2011/837218

[6] Raftari, B. (2013) He's Variational Iteration Method for Nonlinear Jerk Equations: Simple but Effective. *Shock and Vibration*, **20**, 351-356. https://doi.org/10.1155/2013/493048

[7] Rahman, M.S. and Hasan, A. (2018) Modified Harmonic Balance Method for the Solution of Nonlinear Jerk Equations. *Results in Physics*, **8**, 893-897. https://doi.org/10.1016/j.rinp.2018.01.030

[8] Larson, R. and Falvo, D. (2004) Elementary Linear Algebra. 6th Edition.

[9] Lambert, J.D. and Lambert, J. (1973) Computational Methods in Ordinary Differential Equations. Wiley, London, 23.

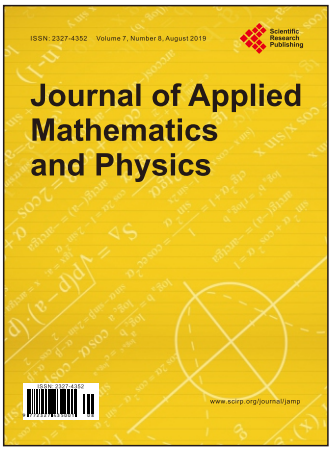
[10] Jator, S.N. (2007) A Sixth Order Linear Multistep Method for the Direct Solution of $y'' = f(x, y, y')$. *International Journal of Pure and Applied Mathematics*, **40**, 457-472.

[11] Dahlquist, G. (1956) Convergence and Stability in the Numerical Integration of Ordinary Differential Equations. *Mathematica Scandinavica*, **4**, 33-53. https://doi.org/10.7146/math.scand.a-10454

[12] Ramos, H., Kalogiratou, Z., Monovasilis, T. and Simos, T. (2016) An Optimized Two-Step Hybrid Block Method for Solving General Second Order Initial-Value Problems. *Numerical Algorithms*, **72**, 1089-1102. https://doi.org/10.1007/s11075-015-0081-8

[13] Mirzabeigy, A. and Yildirim, A. (2014) Approximate Periodic Solution for Nonlinear Jerk Equation as a Third-Order Nonlinear Equation via Modified Differential Transform Method. *Engineering Computations*, **31**, 622-633. https://doi.org/10.1108/EC-02-2012-0024

Call for Papers



Journal of Applied Mathematics and Physics

ISSN Print: 2327-4352 ISSN Online: 2327-4379
http://www.scirp.org/journal/jamp

Journal of Applied Mathematics and Physics is an international journal dedicated to the latest advancement of applied mathematics and physics. The goal of this journal is to provide a platform for researchers and scientists all over the world to promote, share, and discuss various new issues and developments in different areas of applied mathematics and physics. We aim to publish high quality research articles in terms of originality, depth and relevance of content, and particularly welcome contributions of interdisciplinary research on applied mathematics, physics and engineering.

Subject Coverage

The journal publishes original papers including but not limited to the following fields:

- Applications of Systems
- Applied Mathematics
- Applied Non-Linear Physics
- Applied Optics
- Applied Solid State Physics
- Biophysics
- Computational Physics
- Condensed Matter Physics
- Control Theory
- Cryptography
- Differentiable Dynamical Systems
- Engineering and Industrial Physics
- Experimental Mathematics
- Fluid Mechanics
- Fuzzy Optimization
- Geophysics
- Integrable Systems
- Laser Physics
- Methodological Advances
- Multi-Objective Optimization
- Nanoscale Physics
- Nonlinear Partial Differential Equations
- Non-Linear Physics
- Nuclear Physics
- Numerical Computation
- Optical Physics
- Portfolio Selection
- Riemannian Manifolds
- Scientific Computing
- Set-Valued Analysis
- Soliton Theory
- Space Physics
- Symbolic Computation
- Topological Dynamic Systems
- Variational Inequality
- Vector Optimization

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

The journal publishes the highest quality original full articles, communications, notes, reviews, special issues and books, covering both the experimental and theoretical aspects including but not limited to the above materials, techniques and studies. Papers are acceptable provided they report important findings, novel insights or useful techniques within the scope of the journal. All manuscript must be prepared in English, and are subjected to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by prints in hard copy.

Website and E-Mail

http://www.scirp.org/journal/jamp **E-mail: jamp@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: http://www.scirp.org

Subscription: sub@scirp.org

Advertisement: service@scirp.org