OPTIMIZATION OF WEIGHT FUNCTION FOR (3+1)D PHONON PROPAGATION IN WEYL FERMION SEA EXPRESSED BY CLIFFORD ALGEBRA USING ELMAN RNN AND ECHO STATE NETWORK

Sadataka Furui[1] and Serge Dos Santos[2]

[1]Faculty of Science and Engineering, Teikyo University, Utsunomiya, Tochigi, Japan
[2]INSA Centre Val de Loire, Universite de Tours, INSERM, Imaging Brain & Neuropsychiatry iBraiN U1253 F-411034 Blois Cedex, France

ABSTRACT

We modify the lattice simulation of (3+1)D Quantum Chromo Dynamics using fixed point actions by replacing Dirac fermions to Weyl fermions expressed by biquaternions. Paths of phonons are described by the weight function of eigenfunctions. The optimization is performed by using Elman Recurrent Neural Network and the Echo State Networks. Numerical results of the two optimizations are compared. We compare in lower dimensional systems the Time-Reversal based Nonlinear Elastic Wave Spectroscopy and the theory based on Quaternion Field Theory.

KEYWORDS

Clifford Algebra, Quaternion Field Theory, Elman Recurrent Neural Network, Echo State Network

1. INTRODUCTION

In application of Non Destructive Testing (NDT), ultrasonic wave propagation in (3+1) D system including hysteresis effect is important. In current NDT by ultrasonic waves, Time Reversal based Nonlinear Elastic Wave Spectroscopy (TR-NEWS) is applied[1]. In the analysis of experimental data Excitation Symmetry Analysis Method (ESAM) is used, in which quaternion bases are used.

In [2,3,4] we analysed the path of solitonic phonons in neutral Weyl fermion sea described by quaternions. Mathematically, there are 4 division algebras real R, complex C, quaternion H, and octonion O. Quaternions satisfy associativity, and expressed as a I +b I a +cj +d k, a,b,c are real, and ii=jj=kk=-1, ij=-ji, jk=-kj, ki=-ik, ijk=-1.

In 1995, Adler proposed a generalization of Heisenberg picture quantum mechanics using quaternionic Hilbert space [5].

DOI : 10.5121/mathsj.2024.11301

In our analysis of path integral, the fixed point lattice actions for Quantum Chromodynamics (QCD) [6] are modified to consider 27 paths in (3+1)D : A) 7 paths on a 2D plane, B) 13 paths on a 2D plane combined with paths orthogonal to the plane, and C) 7 paths in 3D space combined with paths which have hysteresis effects. In the present work, we describe the C) type paths which include hysteresis effects.

The structure of this presentation is as follows. In Section 2, we show 2 methods of optimising the path. In Section 3, we present numerical results and in section 4 we give conclusion and outlook.

2. OPTIMIZATION OF PATHS

Recently, machine learning (ML) techniques are used in solving various problems of science and engineering [7,8,9]. We performed the optimization of A-type and B-type paths using Python.

Bianchi et al. [10] remarked that for getting the optimized solution in Recurrent Neural Network (RNN), the Elman RNN and the Echo State Network (ESN) are suitable. In this section, we present the two methods.

2.1. Elman Recurrent Neural Network

Elman RNN (ERNN) has the input layer, the hidden layer and the output layer. At each time step t, the input layer processes the component x[t] in R^N_i and a bias b_i in R^N_h . N_i is the number of nodes in the input layer and N_h is that in the hidden layer. They are multiplied by the input weight matrix W_io in R^(N_i x N_h). The internal state of the network h[t-1] and a bias b_h in R^Nh are multiplied by the weight matrix W_hh in R^(N_h x N_h).
The internal state and the output of the time step t satisfy

h[t]=f(W_ih (**x**[t]+**b**_i)+W_hh (**h**[t-1]+**b**_h)
y[t]=g(W_ho (**h**[t]+**b**_o))

The gate function f and g are usually chosen to be logistic sigmoid function or tanh function [11]. In standardization of data of sigmoid function, we used 18 samples for trainings and 8 samples for validations.

In our phonon propagation model, there are 7 paths L19, L20, L21, L22, L23, L24 and L25 that contain time shifts. There are 6 epochs at which more than two paths have time shifts and mixing of the paths occur. The path of L19-L25 are shown in figs 1-7. The balls are the point where quaternion time shifts occur. The ball at right upper corner and at the left lower corner overlap in all paths, and mixing of all 7 paths occur. The ball at the center of L20, L22, L24 and L25 overlap, and mixing of these paths occur.

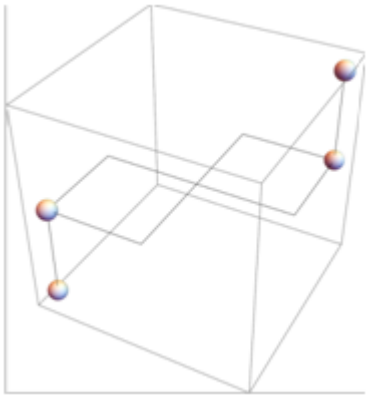

Figure 1. The path of L19.

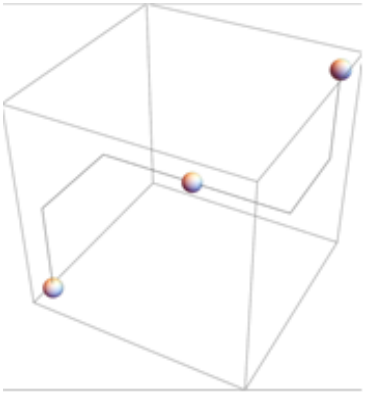

Figure 2. The path of L20

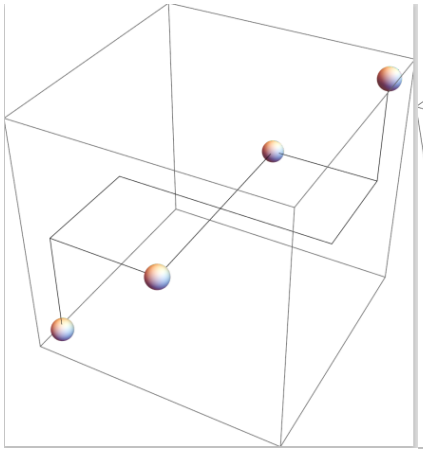

Figure 3. The path of L21.

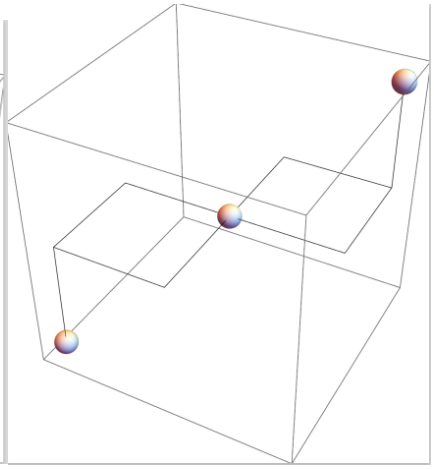

Figure 4. The path of L22

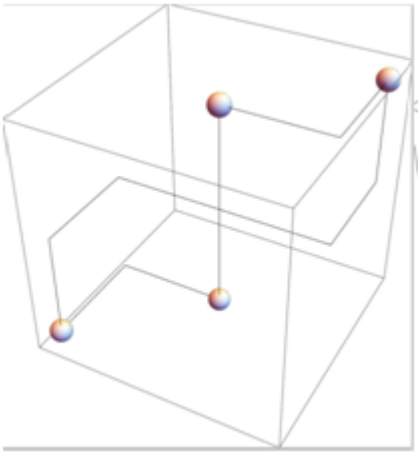

Figure 5. The path of L23.

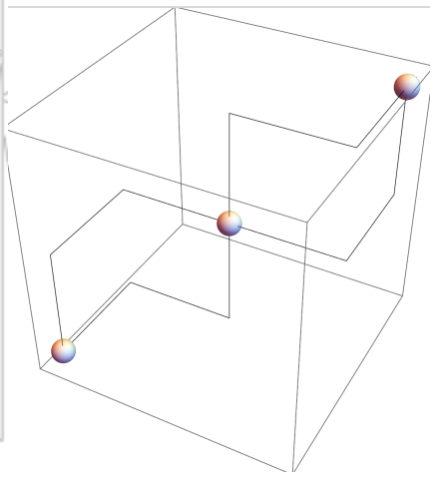

Figure 6. The path of L24

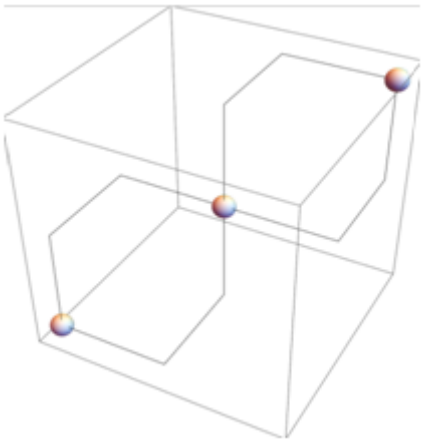

Figure 7. The path of L25.

In biquaternion bases, the quaternion time is represented by e_1e_4, e_2e_4 or e_3e_4. The x coordinate is e_2e_3, the y coordinate is e_3e_1, and the z coordinate is e_1e_2. We choose the pair of e_4 should be contained in the neighboring space coordinate. In the case of L21 and L22, there appear 2 points where both e_1e_4 and e_2e_4 are possible. We distinguish the two by adding L21' and L22' and considered mixing of 9 paths.

We performed ERNN simulation using Mathematica [13]. Since we fixed the W_hh matrix for each cycle and dropped **b**_h, the difference of training data and validation data became negligible after 20 cycles.

2.2. Echo State Network

ESN was employed in reservoir computing framework of chaotic paths [14]. The state update and output are defined as

$$\mathbf{h}[t]=f(W\_rr\ \mathbf{h}[t-1]+W\_ir\ \mathbf{x}[t]+W\_or\ \mathbf{y}[t-1]+\varepsilon),$$
$$\mathbf{y}[t]=g(W\_io\ \mathbf{x}[t]+W\_ro\ \mathbf{h}[t]),$$

where ε is a small noise term. In our case we choose 3D projected momentum space trajectories of L19, L20, L21, L22, L23, L24, L25 given by the fixed-point action as the x[t]. The mixing of L21', L22' are taken into account in the training of the weight matrix W_{rr}. The ESN architecture is shown in Figure 8.
.

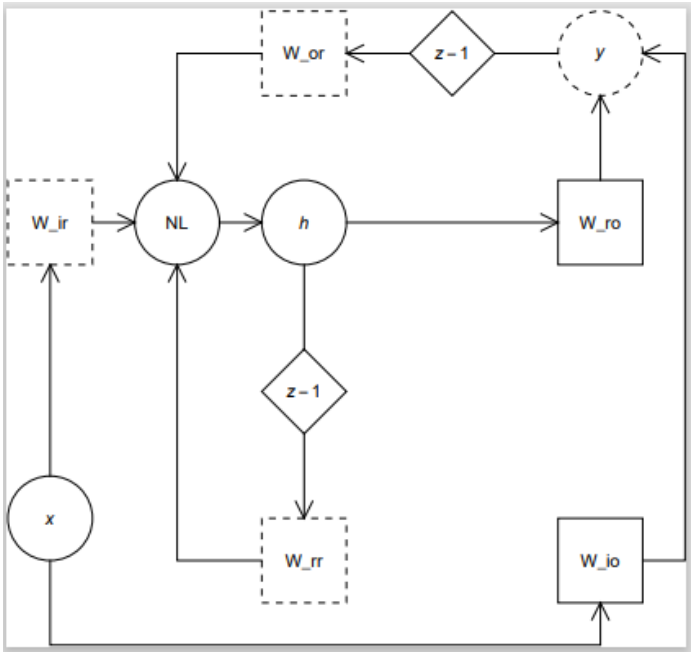

Figure 8. The ESN architecture, modified from Bianchi et al.[10].

The circles represent input x, state h, output y and nonlinear operation NL, respectively. The square W_ro in our model is fixed, W_ir, W_rr, and W_or are randomly initialized trainable matrices, W_io =W_ir W_ro. The z-1 is the unit delay operator.

The weight matrix W_ro is obtained from that used in ERNN. In ERNN L21, and L22 had two components and there were 9 paths, but we replace average of two components and consider 7 paths. W_ir, W_rr and W_or are optimized for minimizing the differences of output y[t] at t=4,5,9,10,11,12 become close to the outputs in ERNN.

We calculate X={**x**[4],**x**[5],…,**x**[16],**x**[1],**x**[2],**x**[3]}, H={**h**[4],**h**[5],…,**h**[16],**h**[1],**h**[2],**h**[3]}, and the tensor S=[X^T, H^T] where X^T means the transpose of X, and the loss function $L=\|S\ W-y^*\|^{2,}$ where W=[W_io, W_ro]. The 16 step makes a cycle and the cycles continue until L becomes small. Since W_io is not fixed, but depends on W_ir which is obtained by learning, Moore-Penrose pseudoinverse method is not applicable for solving the least square problem.

As in ERNN, we introduce a bias vector **b**_h and define

h[t]=f(W_io **x**[t]+W_ro(**h**[t]+**b**_h)),

b_h=-2η(W_ro **h**-**y***)W_ro,

to minimize (W_ro **h**-**y***)^2.

The vectors **x**[t]=x_0 e_0+x_1 e_2e_3+x_2 e_3e_1+x_3 e_1e_2 = (x_1, x_2, x_3, x_0) are in the projected space RP^4. The scale is chosen randomly in the beginning, and multiplied by W_ir to yield 7- dimensional vector **y**[t].

3. NUMERICAL RESULTS

The Figure 9 shows the hysteresis effects observed in ERNN. The action integral at t=4,5,9,10,11 and 12 increase through hysteresis effects.

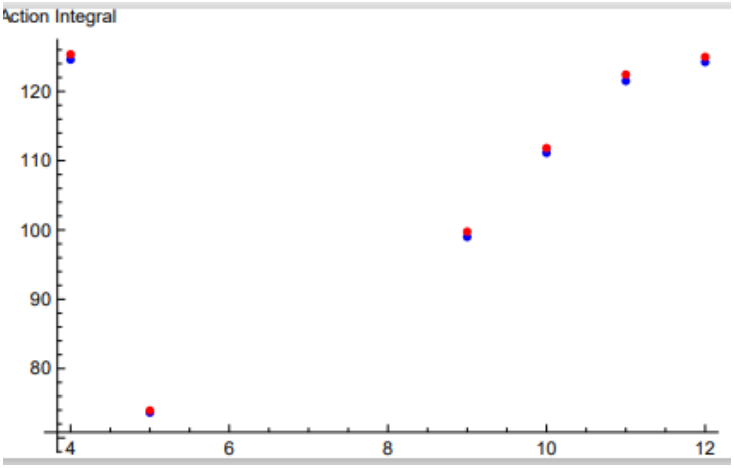

Figure 9. The action integral y^* at t=4, 5, 9, 10,11 and 12 mod 16.. Blue points are results without hysteresis, red points are with hysteresis.

In the ESN, we define the time ta=1,2,3,4,5,6 corresponding to t=4,5,9,10,11,12, respectively. . The mean values and variances of 18 samples are measured for calculations of **h**[t] and **y**[t].

The **h**[t] is a 7-dimensional vector, corresponding to L19, L20, L21, L22, L23, L24 and L25 paths.

The **y**[t] is a 6-dimensional vector, corresponding to outputs of t=4,5,9,10,11 and 12.
The training of W_io, W_ir, W_rr, W_or using Mathematica [13] is not yet finished.

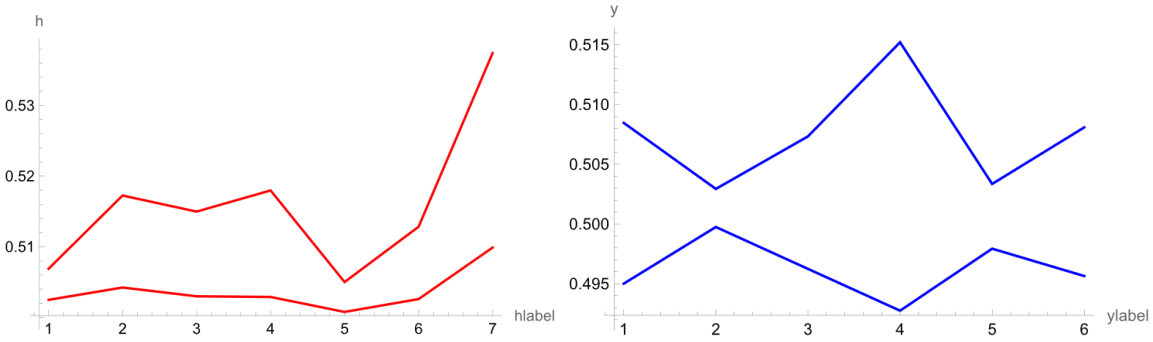

Figure 10. The mean value of **h**[8] and **h**[12] (left) and **y**[8] and **y**[12] (right) as a function of each vector label. **h**[8]> **h**[12]. **y**[8]> **y**[12].

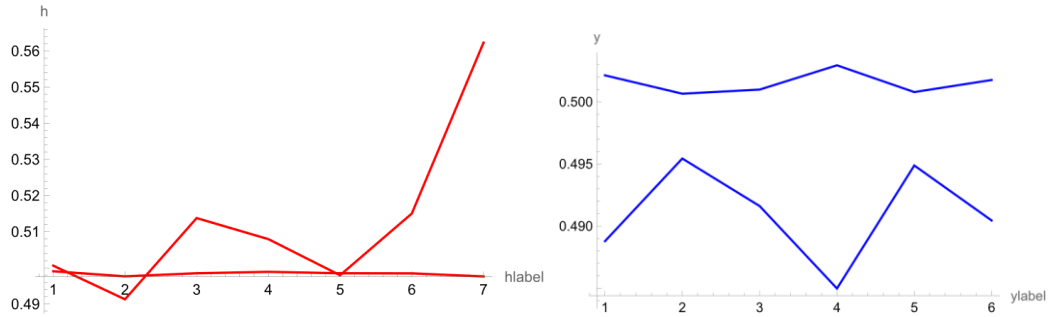

Figure 11. The mean value of **h**[15] and **h**[16] (left) and **y**[15] and **y**[16] (right) as a function of each vector label. **h**[16] has a peak at hlabel=7. **y**[16]> **y**[15]

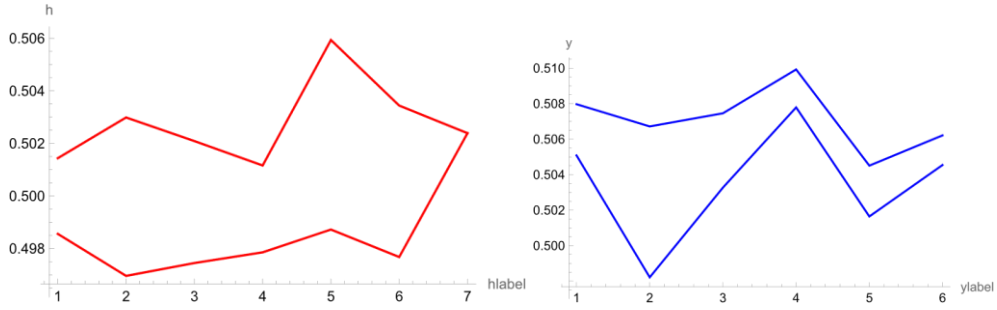

Figure 12. The mean value of **h**[5] and **h**[4'] (left) and **y**[5] and **y**[4'] (right) as a function of each vector label. **h**[5]> **h**[4']. **y**[4']> **y**[5].

h[4'] and **y**[4'] correspond to the beginning of the second cycle. The mean value of **h** of label 7 (L25) at time 5 and at time 4' coincide, while variance of the latter is about twice of the former, which is 0.023.

4. CONCLUSION AND OUTLOOK

The optimization of weight matrices W of ESN are not yet finished. But comparison of the result of ERNN and ESN will be useful to understand the nature of RNN.

We adopted biquaternions for describing the environment in which ultrasonic waves propagate. Quaternions are represented as

q(t)=cos(t/2) + **u** sin(t/2), where $-\pi < t < \pi$ [12] and **u** is fixed by the path of a phonon propagates. The scale of the vector x[t] defines the probability of transition to other paths, or hysteresis effects.

In the standard model, vector $U_V(1)$ symmetry and chiral $U_A(1)$ symmetry are known. In (2+1)D, we observed that chiral U(1) symmetry is broken in time-reversal symmetric system. Time reversal symmetry in (4+1) D is still to be investigated experimentally and theoretically.

ACKNOWLEDGEMENTS

SF thanks the Japan Industrial Science Laboratory (Nissanken) for the financial aid of the travel expense to Blois in November 2023 and Prof. Hamada for allowing using a workstation in his lab.

REFERENCES

[1] Arruga,D., Dos Santos S. and Brochard, M., (2024) "Advanced Excitation Symmetry Analysis Method (ESAM) for improved automated ultrasonic testing of polished materials", e-Journal of Nondestructive Testing- 20th WCNDT (.Incheon Korea, 27-31 May 2024).

[2] Furui,S. and Dos Santos, S., (2023) "Application of Quaternion Neural Network to Time Reversal Based Nonlinear Elastic Wave Spectroscopy", Transactions of Indian National Academy of Engineering, vol.8 (1), pp183-199.

[3] Furui,S. and Dos Santos, S.,(2023) "Clifford Fourier Transforms in (2+1)D Lattice Simulations of Soliton Propagation", PoS Lattice 2022, The 39th International Symposium on Lattice Field Theory (Bonn, Germany, 8-13 August 2022).

[4] Furui,S. and Dos Santos, S.,(2024) "The optimization of paths in the R^{3,1}by Markov Chain Monte Carlos", arXiv:submit/5611595.

[5] Adler, S.L., (1994) "Generalized quantum dynamics", Nuclear Physics B415, pp195-242, North Holland.

[6] DeGrand, T., Hasenfratz, A., Hasenfratz, P. and Niedermayer, F., (1995) "Non-perturbative test of the fixed point action for SU(3) gauge theory " , Nucl. Phys. B454 pp.615-637.

[7] Abbott et al., (2024) "Practical applications of machine -learned flows on gauge fields", PoS Lattice 2023, The 40th International Symposium on Lattice Field Theory,(Fermi National Accelerator Lab..July31-August 4, 2023).

[8] Albergo, M.S. et al.(2021) "Introduction to Normalizing Flows for Lattice Field Theory", arXiv: 210108176 v3[hep-lat]

[9] Bialas, P., Korcyl, P. and Stebel, T. (2022) "Gradient estimators for normalizing flows",, arXiv:2202.01314v1[stat.ML].

[10] Bianchi,F.M. et al., (2017) "Recurrent Neural Networks for Short-Term Load Forecasting",, SpringerBriefs in Computer Science, Springer.

[11] Raschka, S., Liu, Y. and Mirjalili, V., (2022) "Machine Learning with PyTorch and Scikit-Learn", Packt, Birmingham-Mumbai

[12] Gray,A., Abbena, E. and Salamon, S. , (2006) "Modern Differential Geometry of Curves and Surfaces with Mathematica", Chapman & Hall/CRC, Boca Raton..

[13] Wolfram Research (2023) Mathematica 14.

[14] Jaeger, H and Haas, H., (2004) "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication", Science 304 (2. April 2004) Springer.

[15] Bianchi, F.M., Scardapane, S., Uncini, A. ,Rizzi, A. and Sadeghian, A., (2015) "Prediction of telephone calls load using Echo State Network with exogenous variables", Neural Networks 71 pp.204-213, Elsevier.

AUTHORS

Sadataka Furui, Research Collaborator at Faculty of Science and Engineering of Teikyo University, Formerly Professor of Graduate School of Science and Engineering of Teikyo University, Utsunomiya, Japan

Serge Dos Santos, Associate Professor at INSA Centre Val de Loire at Blois, Universite de Tours, INSERM, Imaging Brain & Neuropsychiatry iBraiN U1253 F-411034 Blois, France