Pauli Matrices and the Theory of Representations of the Group of Rotations

Hüseyin ÖNAL

Abstract: : *It is shown that Pauli Matrixes can be derived from irreducible rotation group representations of the weight* $l=\frac{1}{2}$ *, which in turn based on the system of infinitesimal (elementary) spatial rotations. The last permits to substantiate why Pauli matrixes can be so sufficiently used for modeling of physical rotations.*

Keywords: Pauli matrices, Group rotations, spinor transformation, group of rotation.

Introduction

In (Milnikov A.A., Prangishvili A.I., Rodonaia I.D.(2005)- Milnikov A., Onal H., Partskhaladze R., Rodonaia I., (2004)) we get new algorithm based on spinor representation of 3-dimensional rotation group. Pauli matrixes were used essentially but there was not mentioned deep connection between physical rotations and the matrixes. The presented work is devoted to illumination of such connection.

Rotations of a three-dimensional space can be described by complex matrices of second order. In that case, we use the stereographic projection of the sphere onto the plane. Each rotation of a three-dimensional space transforms a point of the sphere to another point of the sphere. This process corresponds to the transformation of points in the plane. The relation between the coordinates of the sphere and those of plane points is given by the formulas

$$\xi = \frac{x}{\frac{1}{2}-z}, \eta = \frac{y}{\frac{1}{2}-z} \qquad (1)$$

Where x, y, z are the coordinates of a point on the sphere, while ξ

Hüseyin ÖNAL *is an associate professor in Faculty of Computer Technologies and Engineering at International Black Sea University, Georgia*

and η are the coordinates of the stereographic projection of this point onto the plane. If we introduce the complex variable $\zeta = \xi + i\eta$, then it turns out that to each rotation there corresponds a linear-fractional transformation of the form

$$\varsigma' = \frac{\alpha\varsigma + \beta}{\gamma\varsigma + \delta} \qquad (2)$$

Where α, β, χ, δ are complex constants. To the linear-fractional transformation (2) we can uniquely assign the transformation matrix of second order

$$\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} \qquad (3)$$

acting in the complex plane. By virtue of the linear-fractional property of transformation (2) we can find that determinant (3) is equal to ±1: for this it is sufficient to multiply both the numerator and the denominator of (2) by

$$\pm \frac{1}{\sqrt{\alpha\delta - \beta\gamma}}$$

(clearly, ς' remains unchanged). The latter means that to each rotation there corresponds the transformation of the complex plane by means of matrix (3) defined to within a sign. If determinant (3) is equal to 1, then we have a proper rotation and matrix (3) is unitary and unimodular. In that case, the converse statement is also true: to any unimodular matrix of form (3) there corresponds a rotation. By the unitary property of matrix (3) we obtain $\alpha\bar{\gamma} + \beta\bar{\delta} = 0, \alpha\bar{\alpha} + \beta\bar{\beta} = 1, \gamma\bar{\gamma} + \delta\bar{\delta} = 1$.

If we add here the condition of unimodularity $\alpha\delta - \gamma\beta = 1$, then we easily have $\delta = \bar{\alpha}, \gamma = -\bar{\beta}$, i.e. for rotations matrix (3) takes the form

$$\begin{vmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{vmatrix} \qquad (4)$$

It is possible to obtain Pauli matrixes in the shape of (4).

Elementary Rotations and Group representations

A finite-dimensional representation of the group of rotations G is said to be given if to each element *g* of the group there corresponds a linear transformation T_g in some linear space L^n. It should be emphasized that we consider the group of rotations in the three-dimensional space though the corresponding linear transformations T_g act in L^n T too. In order that this correspondence be a representation, the following conditions must be fulfilled:

$$T_{g_1}T_{g_2} = T_{g_{11}g_2} \quad \text{and} \quad T_e = E \qquad (5)$$

In finite-dimensional spaces, linear transformations are given by means of matrices and therefore finite-dimensional representations are, as a matter of fact, representations of the group of rotations by means of matrices. The representation is called basic when to each solution there corresponds its matrix in L^3. Another example of the representation can be given as follows: take two arbitrary vectors x and y and form n^2 products of their coordinates (this is done in the general case, while for n = 3 we obviously have $3^2 = 9$). It can be shown that in the process of rotation the products $x^i y^j$ undergo transformation (it is defined by the rotation transformation of the vectors x and y) which satisfies conditions (5). Therefore this representation is defined by 3^2 parameters as different, say, from the basic transformation which is defined by three Euler angles.

Linear transformations of representations (matrices) may have invariant subspaces. A subspace is called invariant with respect to a given representation if is invariant with respect to all transformations of given T_g[1]. A representation is called irreducible if it has no invariant subspaces. The study of representations of rotation groups comes to the study of irreducible representations.

As has been mentioned above, rotations can be given by various systems of parameters. In the role of such parameters let us take three numbers ξ_1,ξ_2,ξ_3 which are the coordinates of the vector directed along the rotation axis, the length of which is equal to the rotation angle. The matrix T_g is the function of these parameters $T_g = T_g(\xi_1,\xi_2,\xi_3)$ and, for $\xi_1=\xi_2=\xi_3 = 0$, we have $T_g(0, 0, 0) = E$. The function $T_g = T_g(\xi_1,\xi_2,\xi_3)$ can be decomposed in a neighborhood of the point $\xi_1=\xi_2=\xi_3 = 0$ as follows:

$$T_g(\xi_1,\xi_2,\xi_3) = E + A_1\xi_1 + A_2\xi_2 + A_3\xi_3 + O(\xi_1,\xi_2,\xi_3), \qquad (6)$$

where $A_i = \dfrac{\partial \mathrm{Tg}(\xi_1,\xi_2,\xi_3)}{\xi_i}\Big|\xi_1 = \xi_2 = \xi_3 = 0 (i = 1,2,3)$ are constant matrices;

[1] It is important to emphasize that the subspace must be simultaneously invariant with respect to all T_g of the considered representation, since it is obvious that for each T_g there always exists its invariant subspace defining the rotation axis: this is a one-dimensional subspace in the case n = 3 or a subspace of dimension greater than 1 (2. section but ? n - 2) in the case $n > 3$.

$O(A1.\xi_1,\xi_2,\xi_3)$ are infinitesimal values of higher orders as compared with $\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$

The matrices A_i have a simple physical meaning: they define infinitesimal rotations about the coordinate axis.

Let us now show that the considered representation is completely defined by these matrices. For this, we consider two rotations $g(t\xi_1, t\xi_2, t\xi_3)$ and $g(s\xi_1,s\xi_2,s\xi_3)$ about the vector (ξ_1,ξ_2,ξ_3). From the definition of this vector it follows that the first rotation is the rotation by the angle $t\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$ while the second rotation – by the angle $s\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$ .The product $g(t\xi_1, t\xi_2, t\xi_3)\cdot g(s\xi_1,s\xi_2,s\xi_3)=$ g$((t+s)\xi_1, (t+s)\xi_2, (t+s)\xi_3)$ is obviously the rotation by the angle

$$(t+s)\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$$

Using the first property of the representation we can also write an analogous equality for matrices that realize this representation

$$T((t+s)\xi_1, (t+s)\xi_2, (t+s)\xi_3) = T(t\xi_1, t\xi_2, t\xi_3)\ \mathrm{T}(s\xi_1, s\xi_2, s\xi_3). \qquad (7)$$

Differentiating both sides of (7) with respect to s and then putting $s = 0$, we have

$$\frac{d}{dt}T(t\xi_1,t\xi_2,t\xi_3) = \frac{t}{dt}T(s\xi_1,s\xi_2,s\xi_3)\big|_{s=0}\cdot T(t\xi_1,t\xi_2,t\xi_3) \qquad (8)$$

From (6) it follows that

$$\frac{d}{dt}T(s\xi_1,s\xi_2,s\xi_3)\big|_{s=0} = A_1\xi_1 + A_2\xi_2 + A_3\xi_3 \qquad (9)$$

By substituting (9) into (8) we obtain a system of linear differential equations that allows us to define the elements of the representation matrix

$$\frac{d}{dt}X(t) = (A_1\xi_1 + A_2\xi_2 + A_3\xi_3)\cdot X(t) \qquad (10)$$

Where $X(t) = T(t\xi_1, t\xi_2, t\xi_3)$.

The initial conditions are $X(0) = T(0, 0, 0) = \mathrm{E}$.[2]

[2] Here we must prove the differentiability of representation matrices with respect to the parameters ξ_1,ξ_2,ξ_3 see (Gelfand I.N.(1966)).

The solution of (10) is obviously

$$X(t) = e^{t(A_1\xi_1 + A_2\xi_2 + A_3\xi_3)} \tag{11}$$

From (11) follows

$$T(\xi_1, \xi_2, \xi_3) = e^{(A_1\xi_1 + A_2\xi_2 + A_3\xi_3)} \tag{12}$$

(12) just shows that the considered representation $T(t\xi_1, t\xi_2, t\xi_3)$ is defined by the matrices A_1, A_2, A_3.

Between the matrices A_1, A_2, A_3 there exist the commutation relations

$$[A_1, A_2] = A_3, [A_2, A_3] = A_1 \text{ and } [A_1, A_3] = A_2 \tag{13}$$

where [A, B] = AB – BA is the commutator of the matrices A and B.

If the representation T is unitary, then matrices A_i are skew-Hermitian: $\overline{A}^T = -A$. This follows from the assumption that the representation T is unitarian and from decomposition (6). If we introduce the matrices $H_i = iA_i$ (i = 1, 2, 3)($i = \sqrt{-1}$), then it is obvious that they are Hermitian and the commutation relations take the form

$$[H_1, H_2] = iH_3, [H_2, H_3] = iH_1 \text{ and } [H_1, H_3] = iH_2. \tag{14}$$

Now the problem of defining all possible representations of rotation groups can be reduced to defining such triples of Hermitian matrices H_i that, firstly, satisfy relations (14) and, secondly, actually give irreducible representations of rotations. Thus we define the matrices of infinitesimal rotations A_i too, which, as a matter of fact, is the final purpose of our investigation. A detailed proof of this problem is given in [1]. We only want to make the following remark. Let us introduce the new matrices

$H_+ = H_1 + iH_2, H_- = H_1 - iH_2, H_3 = H_3$. It turns out that if f is the eigenvector of the matrix H_3 corresponding to some eigenvalue of λ (Since H_3 is Hermitian, all λ 's are real numbers), then the vector $f_1 = H_+ f$ either is equal to zero or is the eigenvector of the matrix H_3 corresponding to the eigenvalue $\lambda + 1$. Analogously, the vector $f_2 = H_- f$ either is equal to zero or is the eigenvector of the matrix H_3 corresponding to the eigenvalue λ-1. This property of eigenvectors f_i of the matrix H_3 allows us to prove that their number is equal to $2l + 1$, where l is a half-integer number, i.e. it can be equal either to an integer number or to a half-integer number. Next it is proved that, in the orthonormalized basis constructed of these vectors, the matrices we want to define have the following form:

$$
\begin{aligned}
A_1 f_m &= -iH_1 f_m = -\frac{i}{2}\sqrt{(l+m+1)(l-m)}\, f_{m+1} - \frac{i}{2}\sqrt{(l+m)(l-m+1)}\, f_{m-1},\\
A_2 f_m &= -iH_2 f_m = -\frac{1}{2}\sqrt{(l+m+1)(l-m)}\, f_{m+1} + \frac{i}{2}\sqrt{(l+m)(l-m+1)}\, f_{m-1},\\
A_3 f_m &= -iH_3 f_m = -imf_m,
\end{aligned}
\qquad (15)
$$

where $m = -l, -l+1, \ldots, l$ and f_m are the orthonormalized eigenvectors of the matrix H_3.

The number l is called the weight of the considered irreducible representation.

From (15) it follows that any irreducible representation is defined by its weight uniquely: elementary rotation matrices A_i are defined from (15), while the rotation matrix T is defined from (12). It can be said that (15) gives a complete solution of the problem on defining all irreducible representations of the group of rotations.

Representations of the weight $l = \frac{1}{2}$

Let us find the matrices A_i corresponding to the representation with weight $l = \frac{1}{2}$. If $l = \frac{1}{2}$ then $m = -\frac{1}{2}, \frac{1}{2}$.

We first define A_1.

For $l = \frac{1}{2}$ and $m = -\frac{1}{2}$, from (15) we have

$$
A_1 f_{-\frac{1}{2}} = 0\, f_{-\frac{1}{2}} - \frac{i}{2} f_{\frac{1}{2}} \qquad (16)
$$

Analogously, for $l = \frac{1}{2}$ and $m = \frac{1}{2}$ we have

$$
A_1 f_{-\frac{1}{2}} = -\frac{i}{2} f_{-\frac{1}{2}} - 0\, f_{\frac{1}{2}} \qquad (17)
$$

Thus the matrix A_1 takes the form

$$A_1 = \begin{vmatrix} 0 & -\frac{i}{2} \\ -\frac{i}{2} & 0 \end{vmatrix} \tag{18}$$

It is likewise easy to obtain the expression for the matrices A_2 and A_3:

$$A_2 = \begin{vmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{vmatrix}, \quad A_3 = \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}. \tag{19}$$

Multiplying the matrices A_1 and A_2 by $2i$, and the matrix A_3 by $-i$, we obtain the matrices

$$\sigma_1 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}; \sigma_2 = \begin{vmatrix} 0 & -i \\ i & 0 \end{vmatrix}; \sigma_3 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} \tag{20}$$

Matrices (20) are the so-called Pauli matrices which underlie the spinor theory and which are usually introduced formally without indicating the source of their origination. The latter circumstance makes it difficult to understand their geometrical meaning. However from the above reasoning it follows that these matrices are generated by matrices of infinitesimal rotations about the coordinate axes.

Let us at once indicate their properties

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1 \text{ and} \tag{21}$$

$$\sigma_1\sigma_2 + \sigma_2\sigma_1 = 0; \sigma_2\sigma_3 + \sigma_3\sigma_2 = 0; \sigma_3\sigma_1 + \sigma_1\sigma_3 = 0. \tag{22}$$

(22) shows that the matrices $\sigma_i (i = 1,2,3)$ are anticommutative. Both properties are easily verified in a straightforward manner.

Multiplying the matrices σ_i by $-i$, we obtain the new matrices h_i,

$$h_1 = \begin{vmatrix} 0 & -i \\ -i & \end{vmatrix}; h_2 = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}; h_3 = \begin{vmatrix} -i & 0 \\ 0 & i \end{vmatrix} \tag{23}$$

Which possess other properties that are also easy to verify:

$$h_1^2 = h_2^2 = h_3^2 = -1;$$

$$h_1h_2 + h_2h_1 = -h_3; h_1h_3 + h_3h_1 = -h_2; h_2h_3 + h_3h_2 = -h_1. \tag{24}$$

The matrices σ_i ,h_i(i=1,2,3) are unitary traceless matrices. Their determinants are equal to –1 for σ_i and +1 for h_i,which, in view of their property to be traceless, determines their characteristic polynomials λ^2-1=0 and λ^2+1=0, respectively. Thus, proceeding from the theory of rotation group representation, we can construct (using the matrices σ_i ,h_i(i=1,2,3) and the unit matrix $\sigma_0 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$) the spinor theory and the Clifford algebra. (note that of late it has often been referred to as geometric algebra).

Reference

1. Gelfand I.N. Lectures in Linear Algebra M., (1966), p. 271. Nauka, Moscow

2. Milnikov A., Onal H., Partskhaladze R., Rodonaia I., (2004) Spinor evaluation of Euler's angles: Jornal of applied mechanics, No2, p.p. 48-53. Tbilisi

3. Milnikov A.A., Prangishvili A.I., Rodonaia I.D.(2005) Spinor Model of Generalized Three-dimensional Rotations: Automation and remote Control, Vol.66, No 6, p.p. 876-872. Moscow