International Journal of Theoretical Physics, Vol. 16, No. 4 (1977), pp. 249–257

An Axiomatic Deduction of the Pauli Spinor Theory

H. M. SCHWARTZ

Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701

Received: 22 *September* 1975

Abstract

Pauli's spinor theory is deduced entirely from a postulate relating to the two-valuedness of the spectrum of the Pauli spin operators, without explicit use of the theory of group representations, or any assumption concerning the angular momentum properties of these operators.

1. *Introduction*

Pauli's theory of the spin of the electron (Pauli, 1927) forms an important part of nonrelativistic quantum mechanics, but its original formulation—as is not infrequently true of pioneering contributions—leaves room for further clarification. This is done concisely by Wigner (1959) and by van der Waerden (1974) in terms of the theory of group representations. In this note an alternative approach to Pauli's theory is developed that employs only elementary quantum-mechanics algebra in addition to a minimal postulate reflecting the results of the Stern–Gerlach experiment. As in the approaches of Wigner and of van der Waerden, the angular momentum character of the Pauli spin operators is not assumed but is deduced; and it is also shown, incidentally, that consistency with the results obtained requires essentially that physical space be three-dimensional.

2. *The Postulate*

We consider an n-dimensional ($n \geqslant 3$) Euclidean vector space and a Hermitian operator $\mathbf{S}(u)$, a quantum-mechanical "observable," defined for every *unit vector* u of this space, and subject to the following condition:

> (I) The operator $\mathbf{S}(u)$ has for every u the same eigenvector space and the same two finite nondegenerate eigenvalues s_1, s_2, without its dependence on u being vacuous.

This journal is copyrighted by Plenum. Each article is available for $7.50 from Plenum Publishing Corporation, 227 West 17th Street, New York, N.Y. 10011.

By proper choice of origin and unit for the quantity represented by the observable **S** we can arrange to have

$$s_1 = +1, \qquad s_2 = -1$$

3. *Derivation of a Preliminary Result*

From postulate (I) and equation (2.1) it follows that the operator $\mathbf{S}(u) \equiv \mathbf{S}$ satisfies the condition $\mathbf{S}^2 = \mathbf{I}$, the unit operator. Therefore, since **S** is also Hermitian, and hence

$$\mathbf{S}^\dagger = \mathbf{S} = \mathbf{S}^{-1}$$

S is also unitary. Combined with the fact stemming again from (I) and (2.1), that every matrix representation S of **S** is traceless, it follows readily that S is of the form

$$\begin{pmatrix} \cos\beta & \exp(i\alpha)\sin\beta \\ \exp(-i\alpha)\sin\beta & -\cos\beta \end{pmatrix}, \qquad \alpha, \beta \text{ real} \tag{3.1}$$

From this is deducible the following pivotal result:

(A) When the vectors u, u' are mutually orthogonal, i.e., when $u \cdot u' = 0$, then the matrix of $\mathbf{S}(u)$ in the eigenbase of $\mathbf{S}(u')$ [i.e., the base formed by the eigenvectors of $\mathbf{S}(u')$] has the form

$$\begin{pmatrix} 0 & \exp(i\alpha) \\ \exp(-i\alpha) & 0 \end{pmatrix}, \qquad \alpha \text{ real} \tag{3.2}$$

Let $\psi_1(u)$, $\psi_2(u)$ be normalized eigenvectors of $\mathbf{S}(u)$ associated with the eigenvalues +1 and −1, respectively; denote $\psi_a(u)$ by ψ_a, $\psi_a(u')$ by ψ_a', $\mathbf{S}(u)$ by **S**, and set

$$(\psi_a, \mathbf{S}\psi_b) \equiv S_{ab}, \qquad (\psi_a', \mathbf{S}\psi_b') \equiv S_{ab}', \qquad (\psi_a, \psi_b') \equiv T_{ab} \tag{3.3}$$

with $a, b = 1, 2$. Using the repeated-index summation convention, one has

$$S_{ab}' = (T_{ca}\psi_c, \mathbf{S}T_{db}\psi_d) = T_{ca}^* T_{db} S_{cd}$$

In particular, since

$$\| S_{ab} \| \equiv S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.4}$$

one finds that

$$S_{11}' = |T_{11}|^2 - |T_{21}|^2, \qquad S_{22}' = |T_{12}|^2 - |T_{22}|^2 \tag{3.5}$$

By the last of equations (3.3), $|T_{ab}|^2$ is the probability of measuring the eigenvalue s_a in the state ψ_b'. Hence S_{11}', S_{22}' represent the differences of the probabilities of measuring the eigenvalues +1 and −1 in the states ψ_1' and ψ_2', respectively. One may have an intuitive feeling arising from symmetry considerations that these two quantities are zero. We proceed to prove that this is indeed the case. To begin with, it is clear that these quantities must have the form of the diagonal elements of the general matrix (3.1). This is verified at once from the unitarity of the matrix T, which shows that $|S'_{11}| \leqslant 1$, so that we can write

$$S'_{11} = \cos\lambda \tag{3.6}$$

By the same unitarity we find the relation $S'_{22} + S'_{11} = 0$, which also agrees, as it must, with the form of the matrix (3.1). To complete the proof of the result (A) it remains thus to show that the quantity (3.6) actually vanishes. This is accomplished upon invoking the isotropy property of the underlying vector space [which implies that the value (3.6) does not depend on the choice of the vector pair u, u', subject to the relation, $u \cdot u' = 0$], by evaluating the first of expressions (3.5) explicitly with respect to two unit vectors u_1, u_2 that are mutually orthogonal and each orthogonal to u. The operators $\mathbf{S}(u_1)$ and $\mathbf{S}(u_2)$ have in the eigenbase of $\mathbf{S}(u)$ the matrix representations (3.1) with $\beta = \lambda$ and $\alpha = \alpha_1$ and α_2, respectively. Their normalized eigenvectors are thus found to have, in understandable notation, the expressions (unique up to an inconsequential phase factor)

$$\psi_{\pm}^{(k)} = (\pm \sin\lambda\, e^{i\alpha_k}[2(1 \mp \cos\lambda)]^{-1/2},\ [(1 \mp \cos\lambda)/2]^{1/2}), \qquad (k = 1, 2) \tag{3.7}$$

where the subscripts refer here to the eigenvalues +1 and −1, respectively. By equations (3.6), (3.5), and (3.3), we must have

$$|(\psi_+^{(1)}, \psi_+^{(2)})|^2 - |(\psi_-^{(1)}, \psi_+^{(2)})|^2 = \cos\lambda$$

Substituting from (3.7) and replacing $\cos\lambda$ by w, the last equation reduces to

$$|(1 + w)\exp[i(\alpha_2 - \alpha_1)] + (1 - w)|^2 - |-(1 - w^2)^{1/2}\exp[i(\alpha_2 - \alpha_1)] + (1 - w^2)^{1/2}|^2 = 4w$$

or

$$w(w - 1) = 0$$

It is easily seen that the solution $w = 1$, for which the matrix (3.1) becomes identical with the matrix (3.4), is unacceptable because it renders $\mathbf{S}(u)$ completely independent of u, which is prohibited by the last condition in postulate (I). Hence $w \equiv \cos\lambda = 0$, and the matrix (3.1) assumes the form (3.2).

4. *Connection with the Dimensionality of Physical Space*

From the result (A) it follows, writing again $\mathbf{S}$ and $\mathbf{S}'$ for $\mathbf{S}(u)$ and $\mathbf{S}(u')$ respectively, that

$$[\mathbf{S}, \mathbf{S}']_+ \equiv \mathbf{S}\mathbf{S}' + \mathbf{S}'\mathbf{S} = 0 \qquad \text{when } u \cdot u' = 0 \tag{4.1}$$

This relation will of course be established if it is shown to hold in some given matrix representation and, by the isotropy argument, for some orthogonal pair u, u'. (It is tacitly assumed throughout that all vectors whose designation involves the symbol u are unit vectors). It suffices then to evaluate the anticommutator of the matrices (3.4) and (3.2), and this does indeed vanish.

The result (4.1) leads to the following basic conclusion:

(B) If u_i $(i = 1, 2, 3)$ form an orthonormal triplet, and α_1, α_2 are the values of the parameter α of (3.2) associated with the matrices of $\mathbf{S}(u_1)$, $\mathbf{S}(u_2)$ in the eigenbase of $\mathbf{S}(u_3)$, then

$$\alpha_2 - \alpha_1 = \pm \pi/2 \qquad (\text{modulo } 2\pi) \tag{4.2}$$

This can be seen with the aid of a formula, that will be needed later in another connection, namely,

$$[M_1, M_2]_+ = 2\cos(\alpha_1 - \alpha_2) \cdot I \tag{4.3}$$

where M_i represents the matrix (3.2) for $\alpha = \alpha_i$ $(i = 1, 2)$. This identity is verified immediately, and the result (4.2) follows by combining it with the result (4.1).

The result (B) shows directly that the dimensionality of our underlying vector space, and hence of our physical space as well, cannot be >3. For if it were, then there would exist in addition to the vectors u_i $(i = 1, 2, 3)$ of theorem (B), a fourth vector u_4 normal to all of them. In that case, retaining the eigenbasis of u_3, we would have simultaneously the pair of relations $\alpha_4 - \alpha_1 = \pm \pi/2$ and $\alpha_4 - \alpha_2 = \pm \pi/2$, which is obviously inconsistent with the relation (4.2).

5. *General Determination of the Operators S(u)*

Considering first a right-handed triplet of directions u_i $(i = 1, 2, 3)$ in our physical space, and working in the eigenbase of $\mathbf{S}(u_3)$, the matrix $S(u_3)$ is given by equation (3.4), while some arbitrariness remains in fixing $S(u_1)$ and $S(u_2)$ subject to the form (3.2) and the result (B). In the first place, owing to the arbitrariness in a phase factor of a normalized state vector, there is an associated arbitrariness in the off-diagonal matrix elements of S; namely, S_{12} and S_{21} can be multiplied by factors of the form $\exp(i\sigma)$ and $\exp(-i\sigma)$, respectively (σ real). It constitutes, therefore, a permissible normalization to choose $\alpha = 0$ in (3.2) as corresponding, say, to the direction u_1 which we shall in the sequel take as the direction of the x-axis of a right-handed rectangular coordinate frame. In the second place, there is freedom in the choice of the sign in equation (4.2). Choosing the minus sign conforms with the original convention of Pauli, and our matrices $S(u_i)$ indeed coincide with the *Pauli matrices*:

$$S(u_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv \sigma_x, \qquad S(u_2) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \equiv \sigma_y, \qquad S(u_3) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \sigma_z \tag{5.1}$$

It remains to find the expression for the matrix $S(u)$ for an arbitrary u, in the eigenbasis of $\mathbf{S}(u_3)$, and under the normalization represented by (5.1). This can be done in two steps as follows:

If u is normal to u_3, let ϕ denote the angle from u_1 to u ($u_1 \to u_2$ determining the positive sense of rotation), and set

$$\alpha = f(\phi) \qquad (0 \leqslant \phi < 2\pi) \tag{5.2}$$

By the normalization which we have introduced,

$$f(0) = 0 \tag{5.3}$$

From the latter equation and equation (4.3), we have, in understandable notation,

$$[M, M_1]_+ = 2 \cos f(\phi) \cdot I \tag{5.4}$$

Under the rotation $\phi \to \phi + \phi' \equiv \phi''$, equation (5.4) is transformed into

$$[M'', M']_+ = 2 \cos\, [f(\phi'') - f(\phi')] \cdot I$$

which is independent of the choice of phase normalization, and thus, in view of the isotropy of space, must yield the same value as expression (5.4), so that

$$\cos\, [f(\phi'') - f(\phi')] = \cos f(\phi) \tag{5.5}$$

Taking account of (5.3), there follows the identity

$$f(\phi'') - f(\phi') = \pm f(\phi) \tag{5.6}$$

The arbitrariness in sign can be disposed of by considering the sister identity to (4.3), obtained by taking the corresponding commutator, namely,

$$[M_1, M_2] = 2i \sin(\alpha_1 - \alpha_2)\, \sigma_z \tag{5.7}$$

By reasoning similar to that which led to (5.5), one obtains the identity

$$\sin\, [f(\phi'') - f(\phi')] = \sin f(\phi)$$

which rules out the minus sign in (5.6). The resulting identity

$$f(\phi + \phi') = f(\phi) + f(\phi') \tag{5.8}$$

is familiar, and since $f(\phi)$, considered as a real-valued function for $-\infty < \phi < \infty$, must be bounded over every finite interval (being finite for every finite ϕ), it has the solution

$$f(\phi) = C\phi$$

(see Appendix). This is consistent, as it must be, with the normalization represented by the first matrix of the Pauli set (5.1). For agreement with the second matrix, one has to take $C = -1$. Thus, with the Pauli normalization, the matrix (3.2) takes on the form

$$\begin{pmatrix} 0 & \exp(-i\phi) \\ \exp(i\phi) & 0 \end{pmatrix} \equiv \cos\phi \cdot \sigma_x + \sin\phi \cdot \sigma_y$$

In other words, if $u \equiv u(\phi)$ is normal to u_3, then in the eigenbasis of $\mathbf{S}(u_3)$:

$$S[u(\phi)] \equiv S(\cos\phi \cdot u_1 + \sin\phi \cdot u_2) = \cos\phi S(u_1) + \sin\phi S(u_2)$$

Since this relation is valid in any state-vector basis, it holds also if the matrices S are replaced by the operators $\mathbf{S}$. Moreover, the orthogonal pair u_1, u_2 is arbitrary, in view of the isotropy of space. Therefore generally,

$$\mathbf{S}(\cos\phi u_1 + \sin\phi u_2) = \cos\phi\,\mathbf{S}(u_1) + \sin\phi\,\mathbf{S}(u_2), \qquad u_1 \cdot u_2 = 0 \quad (5.9)$$

If u is not normal to u_3, then there exist angles ϕ, θ such that

$$u = \cos\theta \cdot u_3 + \sin\theta \cdot u', \qquad u' = \cos\phi \cdot u_1 + \sin\phi \cdot u_2$$

and an application of equation (5.9) yields the relation

$$\mathbf{S}(u) = \cos\theta\,\mathbf{S}(u_3) + \sin\theta\,\mathbf{S}(u') = \cos\theta\,\mathbf{S}(u_3) + \sin\theta\,[\cos\phi\,\mathbf{S}(u_1) + \sin\phi\,\mathbf{S}(u_2)]$$

Written more compactly, our general result is as follows (summation over repeated indices $i, j, \ldots$ over the range 1, 2, 3 being assumed here and in the sequel):

(C) If (u_1, u_2, u_3) is an orthonormal vector triplet then

$$\mathbf{S}(a_i u_i) = a_i\,\mathbf{S}(u_i); \qquad a_i \text{ real, and } a_i a_i = 1 \quad (5.10)$$

6. *The Transformation of Pauli Spinors under Rotations*

In the light of the preceding results it is clear that the eigenvector $\psi \equiv \psi(u)$ of the operator $\mathbf{S} \equiv \mathbf{S}(u)$ given in equation (5.10), and corresponding in our units to the eigenvalue +1, can be identified with a general *Pauli spinor*, the operator $\mathbf{S}$ itself representing the projection of the intrinsic angular momentum of an elementary particle of spin 1/2 along the direction given by u. The transformation of ψ under a rotation of axes $x \to x'$, associated with the proper orthogonal matrix R is therefore of the form

$$\psi \overset{R}{\to} \psi' = \mathbf{U}\psi \quad (6.1)$$

where the corresponding transformation of $\mathbf{S}$ is

$$\mathbf{S} \overset{R}{\to} \mathbf{S}' = \mathbf{U}\mathbf{S}\mathbf{U}^\dagger \quad (6.2)$$

Our task is then to determine the unitary operator $\mathbf{U}$, which satisfies equation (6.2) for any given u, and by equation (5.10), it suffices to solve this problem when $\mathbf{S} = \mathbf{S}(u_i) \equiv \mathbf{S}_i (i = 1, 2, 3)$.

First, we note that a rotation of axes

$$x \to x' : x'_i = R_{ij} x_j$$

is represented by the associated transformation

$$u_i \to u'_i = R_{ij} u_j$$

Since $R_{ij}R_{kj} = \delta_{ik}$, it follows from equation (5.10) that

$$\mathbf{S}(u_i') \equiv \mathbf{S}_i' = R_{ij}\mathbf{S}(u_j) \equiv R_{ij}\mathbf{S}_j$$

Our problem reduces thus, upon taking account of equation (6.2), to finding, if possible, a unitary operator $\mathbf{U}$ which satisfies the following set of equations:

$$R_{ij}\mathbf{S}_j = \mathbf{U}\mathbf{S}_i\mathbf{U}^\dagger (i = 1, 2, 3) \tag{6.3}$$

The matrix U of $\mathbf{U}$ is of course determined only up to a phase factor [which is consistent with the similar arbitrariness in ψ occurring in equation (6.1)], and this can be fixed to make the determinant of U unity. We have thus the familiar result,

$$U = \pm\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \qquad |a|^2 + |b|^2 = 1 \tag{6.4}$$

In the matrix representation (5.1), we have then the equations $R_{ij}\sigma_j = U\sigma_i U^\dagger$, where $\sigma_1 \equiv \sigma_x$, etc., or

$$R_{ij}\sigma_j U = U\sigma_i (i = 1, 2, 3) \tag{6.5}$$

We solve these equations by introducing as parameters the Euler angles ϕ, θ, ψ (as defined in Goldstein, 1950). For the rotation about the z-axis with angle ϕ, the equation for $i = 3$ in (6.5), namely the vanishing of the commutator of U and σ_3, yields immediately, in view of equation (6.4), the result $b = 0$; and the equation for $i = 1$ or $i = 2$, the result $a = \pm\exp(-i\phi/2)$. For a rotation about the x-axis with the angle θ, the first equation (6.5), namely the vanishing of the commutator $[\sigma_1, U]$, one finds that $a = a^*, b = -b^* \equiv iB$; then from either the second or the third equation (6.5), it follows that

$$B/a = -\sin\theta/(\cos\theta + 1) = -\tan\theta/2 \tag{6.6}$$

When this equation is combined with the second equation in (6.4), we get $a = \pm\cos\theta/2$, and applying equation (6.6), $B = \mp\sin\theta/2$. Consequently,

$$U = \pm\begin{pmatrix} \exp(-i\psi/2) & 0 \\ 0 & \exp(i\psi/2) \end{pmatrix} \cdot \begin{pmatrix} \cos\theta/2 & -i\sin\theta/2 \\ -i\sin\theta/2 & \cos\theta/2 \end{pmatrix} \cdot$$

$$\cdot \begin{pmatrix} \exp(-i\phi/2) & 0 \\ 0 & \exp(i\phi/2) \end{pmatrix}$$

$$= \pm\begin{pmatrix} \cos(\theta/2)\exp[-i(\phi+\psi)/2] & -i\sin(\theta/2)\exp[i(\phi-\psi)/2] \\ -i\sin(\theta/2)\exp[i(\psi-\phi)/2] & \cos(\theta/2)\exp[i(\phi+\psi)/2] \end{pmatrix},$$

which agrees with the corresponding result on page 613 of Pauli's paper (1927), upon correcting an obvious misprint, and taking account of the difference in sign conventions in Goldstein, 1950 and in Klein and Sommerfeld, 1897 (to which Pauli refers).

7. *Summary*

Pauli's theory of the spin of an electron, as first enunciated by him in the introductory part of his fundamental paper (1927), is developed here axiomatically, using essentially only Postulate I and the isotropy property of the Euclidean metric of physical space. Because of our knowledge of the structure of the two-dimensional linear group representation of the rotation group of our physical space, the results presented in the present paper are in themselves not surprising. They are nevertheless illuminating in bringing out explicitly the intimate connection between the quantum-mechanical intrinsic angular momentum properties of our nontransitory elementary particles and the basic properties of our physical space. The analysis employed represents incidentally an alternative derivation of the fundamental representation of the three-dimensional rotation group.

In addition to the classic work of Wigner and of van der Waerden referred to in Section 1, I should mention Feynman's interesting and instructive discussion of the subject in his Lectures (1965). This was called to my attention after writing of this paper was completed without my being acquainted with that work, with the comment that there exists significant similarity in the two treatments. Upon examining it, however, it seems to me that the latter comment overlooks nontrivial differences in the methods employed.

Appendix

Suppose that $f(x)$ is a real-valued function of x for $-\infty < x < \infty$, satisfying the identity

$$f(x+y) = f(x) + f(y) \tag{A1}$$

and the inequality,

$$|f(x)| < A < \infty \text{ in a bounded interval } I \tag{A2}$$

By well-known simple steps it can be shown that the function $g(x) \equiv f(x) - Cx$ vanishes for all rational x, where $C = f(1) \neq \infty$. Since $g(x)$ satisfies the identity (A1), and by condition (A2), $|g(x)| < B < \infty$ in I, the existence of a real number ξ such that $g(\xi) = K \neq 0$ would lead to an absurdity: For a positive integer n such that $nK > B$, and a rational number y for which $n\xi + y$ is in I, we would have

$$B > |g(n\xi + y)| = |g(n\xi)| = n\,|g(\xi)| = nK > B$$

Hence we must have $g(x) \equiv 0$, i.e., $f(x) \equiv Cx$.

Acknowledgment

I wish to thank my colleague, Professor M. Lieber, for reading this paper.

References

Feynman, R. P., Leighton, R. B., and Sands, M. (1965). *The Feynman Lectures on Physics*, Vol. III (Addison-Wesley, Reading), Chap. 6.

Goldstein, H. (1950). *Classical Mechanics* (Addison-Wesley, Reading), Sec. 4–4.

Klein, F., and Sommerfeld, A. (1897). *Theorie des Kreisels*, Vol. I (Teubner, Leipzig), Sec. 2.

Pauli, W. (1927). *Z. Physik*, **43**, 601.

van der Waerden, B. L. (1974). *Group Theory and Quantum Mechanics* (Springer-Verlag, New York, Heidelberg, Berlin), Chap. IV.

Wigner, E. P. (1959). *Group Theory* (Academic Press, New York and London), Chap. 20.