OPEN ACCESS
IOP Publishing
European Journal of Physics
Eur. J. Phys. 44 (2023) 045401 (13pp)
https://doi.org/10.1088/1361-6404/acdb0e

# Dirac equation from stereographic projection of the momentum sphere

**Michele Marrocco**

ENEA (Italian National Agency for New Technologies, Energy and Sustainable Economic Development), via Aguillarese 301, Rome I-00123, Italy

E-mail: michele.marrocco@enea.it

Received 27 January 2023, revised 24 April 2023
Accepted for publication 1 June 2023
Published 29 June 2023

**Abstract**

The Dirac equation is commonly demonstrated under stringent hypotheses and after considerable math work made in relativistic quantum mechanics and quantum field theory. Here, a purely geometric approach free from hypotheses is suggested. The suggestion draws inspiration from the technique of stereographic projection that was developed before the quantum era to solve gyroscopic problems of classical mechanics. The projected variable is the generalized (or canonical) momentum vector. Its undetermined geometric orientations define a sphere in the momentum space and the projection onto the equatorial plane generates the Pauli matrices as soon as the conventional stereographic matrix is introduced. The associated eigenvalue problem results in the Dirac equation and the eigenvector (bispinor) has components that are related to geometric elements of the momentum space. The procedure has the advantage of revealing the correct form of the Dirac matrices without the mathematical effort that characterizes the presentation of the equation in traditional approaches. The other remarkable advantage is that, unlike the common reduction to the case of free space, the spatial inhomogeneity due to interaction potentials is included in the demonstration from the very beginning. The whole suggestion has an interdisciplinary character (relativity, complex analysis, rotation of rigid bodies, Pauli matrices) and can be useful in teaching the equation to students who lack in sufficient knowledge of quantum mechanics. Students equipped with more advanced education can benefit from the purely geometric perspective of this work if used to complement their studies about the equation.

Supplementary material for this article is available online


0143-0807/23/045401+13$33.00 Printed in the UK

Keywords: Dirac equation, Pauli matrices, stereographic projection, rotation of rigid bodies

(Some figures may appear in colour only in the online journal)

## 1. Introduction

The Dirac equation has a fundamental role in our understanding of quantum electrodynamics and marks the birth of relativistic quantum mechanics (RQM) [1, 2] and quantum field theory (QFT) [3–5]. Despite the common subject, the equation is taught according to different criteria in RQM and QFT courses.

The RQM derivation does not differ from the discovery [6]. It relies on the invariant energy-momentum relationship and the preservation of the relativistic symmetry between spatial and time variables [1, 2]. Importantly, the equation is obtained for free fields. They guarantee the isotropy and the space homogeneity necessary to make the educated guess of constant elements in the matrices that generate the Dirac equation.

A second and more elaborated approach to the equation is typical of QFT courses whose focus is on the transformation laws of fields [3–5]. The approach digs deep into abstract algebra for the purpose of finding the generators of the Lorentz group that dictates the construction of a proper covariant Lagrangian obeying the variational principle.

It is fair to recall that, over the years, alternative attempts have been suggested. They have produced original perspectives that differ in some respects from those characterizing RQM and QFT derivations [7–17]. Despite the interesting initiatives to bring out other physical meanings behind the Dirac equation, none of the alternative proposals has the necessary didactic strength to compete with the two conventional approaches (an explanation for their minor educational potential is available in the supplementary material).

Here, we illustrate a very simple procedure that adds a geometric perspective to the abstract math used in RQM and QFT approaches. The procedure relies on the visual interpretation of the momentum thanks to the technique of stereographic projection so useful in handling gyroscopic problems in classical mechanics of rigid bodies [18, 19]. Among them, the unit sphere sets a paradigmatic example. Its motion with reference to its equatorial plane is fundamental for the formal characterization of the stereographic projection in relation to the Möbius transformation introduced after the extension of the equatorial plane to complex numbers [20]. The transformation connects two points of the complex equatorial plane through a linear fractional relationship established for the motion of the projected point taken on the unit sphere (also named Riemann sphere when referred to complex coordinates). This operation, mediated by the stereographic projection, has profound implications related to the conformal (e.g., angle-preserving) properties that characterize the transformation between the complex plane and the non-Euclidean geometries of the sphere and pseudosphere [20].

Astonishingly, the stereographic projection and the associated Möbius transformation are much more than captivating tools of complex analysis. Physically, they play a significant role in special relativity thanks to a direct connection with the Lorentz transformations [21]. Adding to the interdisciplinary value of the stereographic projection, we must underline that the technique introduces spin matrices and spinor fields in a natural way as a part of the representation process of real three-dimensional (3D) vectors subject to rotations which, by the way, can be viewed as one family of the Lorentz transformations. Given this contiguity with Lorentz invariance, it is intuitive to envision a simple application of the projection to the derivation of the Dirac equation which, in essence, is the extension of special relativity to spinor fields. The extension entails, actually, a reduction in the dimensionality of the problem. In agreement with the meaning of the term, the stereographic projection determines the solution of the 3D dynamical problem in the two-dimensional (2D) vector space made of spinors. The shift is not less informative than solving the equations of the motion in the 3D setting typical of three-vectors of Hamiltonian and Lagrangian mechanics. In place of equations of motion, care is here taken to determine the motion that, when viewed by different inertial observers, guarantees the invariance of the relativistic interval of four vectors undergoing Lorentz transformations. We will learn that the determination of the relativistic motion is less problematic if we apply the stereographic projection.

Common aspects and differences between the suggested approach and conventional demonstrations of the Dirac equation are described in the text. The comparison shows that the current attempt is simpler, allows for the presence of external fields from the very beginning and can be used to introduce the Dirac equation to students with an elementary knowledge of quantum mechanics. Those students that have studied the equation within RQM and QFT contexts can consolidate their studies with the visual approach of the current proposal.

The work is organized as follows. In section 2, a brief summary of RQM and QFT approaches is presented. The notions of the momentum sphere and its stereographic projection are introduced in section 3. This preliminary material is used in section 4 where the generalized (or canonical) momentum is still a vector and the classical version of the Dirac equation is obtained. In section 5, comparisons with conventional approaches are made. The conclusions are in section 6.

## 2. Conventional approaches to the Dirac equation

The Dirac equation has been the subject of several investigations of general pedagogic character [22–28] and many are the applicative examples of didactic interest (a couple of examples are: relativistic particle in a box [29] and linear potential [30]). In these works, the equation is treated under either the RQM recipe or the QFT argument. Notably, differences between the two approaches are a reflex of the distinct viewpoints characterizing RQM and QFT contexts. The former is more analytical, the latter capitalizing on the imaginative power of abstract algebra and group theory. In the following, such differences are briefly elucidated.

### 2.1. RQM derivation

The first approach summarized here is unchanged since its introduction in 1928 [6]. Besides the extensive demonstration of common RQM textbooks [1, 2], short reviews are available [26]. In brief, the criteria used in the derivation are: (i) Lorentz invariance, (ii) dependence on first-order time derivative (iii) analogous dependence for spatial derivatives (iv) limitation to the free field.

Application of the above-mentioned criteria begins with the second. It means that the Dirac equation is introduced *à la* Schrödinger

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = H\psi, \tag{1}$$

where, as anticipated, the first-order time derivative replicates the structure of the time-dependent Schrödinger equation. Adherence to the Lorentz invariance comes with the energy-momentum relationship $E^2 = c^2p^2 + (mc^2)^2$ rewritten as the Klein–Gordon equation after the usual quantum-mechanical replacements $E \to \mathrm{i}\hbar\partial/\partial t$ and $\boldsymbol{p} \to -\mathrm{i}\hbar\boldsymbol{\nabla}$. Squaring equation (1) and equating the result to the Klein–Gordon equation results in the appearance of the Pauli vector $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ whose components are the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2}$$

Finally, the Dirac Hamiltonian is found

$$H_\mathrm{D} = c\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2, \tag{3}$$

where again $\boldsymbol{p}$ is the quantum-mechanical momentum and

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{4}$$

are $2 \times 2$ block representations of $4 \times 4$ matrices containing respectively the Pauli matrices and the identity matrix $I$. Consequently, interpreting equation (1) in matrix form, $\psi$ has to be considered as a multi-component quantity (i.e. bispinor). Note that the physical model is free from interaction terms. This choice may appear arbitrary and uninteresting because we look for an equation that explains how physics works under interaction. Nonetheless, the choice of free fields is instrumental in achieving the main objective. Removal of any possible unbalance in the field evolution across space (as a result of particular space-dependent interactions) helps in assuming the crucial independence of $\boldsymbol{\alpha}$ and $\beta$ on spatial coordinates. In other terms, the coefficients in equation (3) (i.e. matrix elements of $\boldsymbol{\alpha}$ and $\beta$) must contain numerical constants only.

### *2.2. QFT derivation*

A more sophisticated derivation of the Dirac equation revolves around the representation of the Lorentz group [3–5]. A short review of the QFT derivation is available together with the description of some drawbacks [27]. Here, only a very brief summary is given. Considering the advanced level of knowledge necessary to properly capture the key points of the argument, the following short outline can only have a modest descriptive value suitable for readers of limited familiarity with the abstract algebra of the QFT derivation. For this reason, the only purpose of this subsection is to succinctly rough out the criteria guiding the QFT search so that the contrast with the derivation suggested in this work may be appreciated. Those readers interested in more solid details about the QFT derivation are referred to the proper educational literature [3–5].

The whole procedure consists in pursuing the following resolutions: (i) search for a suitable representation of the Lorentz group, (ii) choice of linear transformations for multi-component fields, (iii) definition of a Lagrangian that, thanks to elementary bilinear forms, is a Lorentz scalar, (iv) application of the variational principle.

The QFT derivation begins with (i) Suppose to introduce a Lorentz operator $\Lambda$ that is linear and transforms the vector $x$ in a new vector $x'$. In compact form, the transformation is $x' = \Lambda x$. The transformation is expected to induce some change in the field that is assumed to be made of a multiplet of dimension $n$. The change might be very subtle but, requiring linearity, can be represented by means of $n \times n$ matrix $M(\Lambda)$ that depends on the specific Lorentz operator $\Lambda$. Therefore, the following linear transformation is hypothesized

$$\Phi_i(x) \to \Phi'_i(x') = M_{i,j}(\Lambda)\Phi_j(\Lambda^{-1}x'), \tag{5}$$

where summation is meant for repeated indices. The right side of equation (5) says that the $i$th component of the field evaluated at the transformed point transforms with the matrix $M(\Lambda)$ applied to the components of the field evaluated at the point before the transformation. The transformation law can be condensed in $\Phi \to M(\Lambda)\Phi$ and acceptable forms of matrices $M$

are restricted to those satisfying the fundamental group property $M(\Lambda) = M(\Lambda')M(\Lambda'')$ when $\Lambda = \Lambda'\Lambda''$. This property translates the fact that two subsequent Lorentz operators $\Lambda'$ and $\Lambda''$ are completely equivalent to the Lorentz operator $\Lambda$ resulting from the combination of $\Lambda'$ and $\Lambda''$. The problem is now to find the possible $n$-dimensional matrices $M$ that form the so-called representation of the Lorentz group.

The illustrative example of rotations in real space is very useful to build an approach valid for the Minkowski space where, eventually, properties of the Lorentz group refer to. Considering the Lie group (continuous group infinitesimally close to the identity), the generators of the group are the angular momentum operators. When these operators are written as an anti-symmetric tensor for the Minkowski space, they are

$$J^{\mu,\nu} = \mathrm{i}(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu}), \tag{6}$$

(the typical QFT convention $\hbar = c = 1$ is used) and it is possible to show that these operators generate three boosts and three rotations. Based on the commutation relations of the generators of the group, the simplest representation of the Clifford algebra in terms of $4 \times 4$ matrices can be defined. The definition leads to the so-called Weyl or chiral representation of the Dirac matrices and this introductory material is finally used to derive the bilinear forms for the free field. Among them, the most basic forms are picked up and, with the benefit of hindsight, the Lorentz-invariant Dirac Lagrangian for the free field is hypothesized

$$L = \overline{\psi}(\mathrm{i}\gamma^{\mu}\partial_{\mu} - m)\psi, \tag{7}$$

where $\bar{\psi} = \psi^{\dagger}\gamma^{0}$ and $\gamma^{\mu}$ are the Dirac matrices. The variational principle applied to $L$, results in $(\mathrm{i}\gamma^{\mu}\partial_{\mu} - m)\psi = 0$, which agrees with equations (1) and (3) if the fundamental constants $\hbar$ and $c$ are reinstated in their proper places and the connection of the Dirac matrices with $\boldsymbol{\alpha}$ and $\beta$ is recalled.

## 3. Stereographic projection of the momentum sphere

To fulfill the objective of a simple and sound introduction to the Dirac equation, we use the geometric construct of the stereographic projection of the momentum. To that aim, we need some preparatory material and, for this reason, the current section introduces the concept of the momentum sphere and its stereographic representation that will be used in the following section to accomplish the derivation of the classical Dirac equation.

### *3.1. Momentum sphere*

In the descriptive summaries of section 2, the attentive reader should have noticed that the RQM and QFT approaches rely on the free-space limit where the momentum is a constant of the motion. Despite the pedagogic interest in such a basic circumstance, the primary importance of the Dirac equation comes from the presence of external fields and, indeed, one convincing proof of the equation is its successful application to electromagnetic interactions. For this reason, let us take the general case occurring when the momentum undergoes changes due to some interaction $V$ at position $\boldsymbol{r}$ where the mass $m$ is localized. It means that attaching the meaning of generalized momentum to $\boldsymbol{p}$, relativity is stipulated in the relationship

$$[E - V(\boldsymbol{r})]^2 = c^2p^2 + (mc^2)^2. \tag{8}$$

We recall that the generalized momentum results from its canonical definition as the derivative of the Lagrangian [18, 19] and is made of the actual momentum $\boldsymbol{p}_0$ plus the term

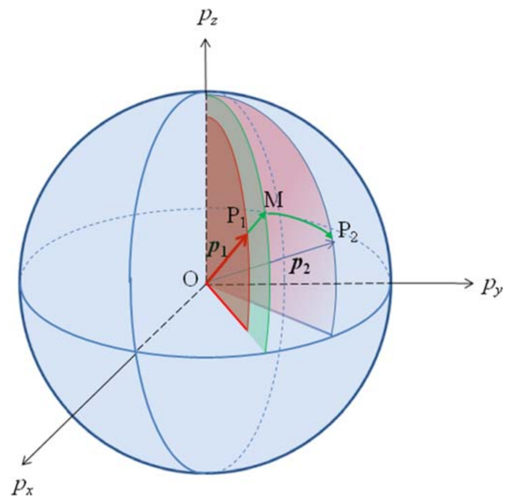


**Figure 1.** Representation in the momentum space of the evolution from the initial generalized momentum $\boldsymbol{p}_1$ (red arrow) to the final generalized momentum $\boldsymbol{p}_2$ (light blue arrow). The points $M$ and $P_2$ lie on the same spherical surface and are connected by a geodesic (green arc of the circle going through $M$ and $P_2$).

proportional to the vector potential. However, by letting $V = 0$ and $\boldsymbol{p} = \boldsymbol{p}_0$ in equation (8), we recover the free-field case treated in RQM and QFT contexts.

The first thing to note about equation (8) is that the absolute value of the generalized momentum is a function of the space coordinates through the scalar potential. It means that taking two time instants $t_1$ and $t_2$ we expect the change from the initial generalized momentum $\boldsymbol{p}_1$ to the final generalized momentum $\boldsymbol{p}_2$. Thus, the knowledge of the scalar potential at the coordinates $\boldsymbol{r}_1 = \boldsymbol{r}(t_1)$ and $\boldsymbol{r}_2 = \boldsymbol{r}(t_2)$ gives information about the absolute values $p_1$ and $p_2$, but nothing is known about the direction of the final vector $\boldsymbol{p}_2$. This situation is depicted in figure 1, where the frame of reference is taken in the momentum space. Note that the initial vector $\boldsymbol{p}_1$ has an arbitrary direction and length $\mathrm{OP}_1 = p_1$, whereas the final momentum vector $\boldsymbol{p}_2$ has a different direction and length $\mathrm{OP}_2 = p_2$. In this example, $p_2 > p_1$ but one could take the inverse relationship without loss of generality.

Although the specific choice of $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ in figure 1 is made for graphical purposes, it is intuitive that the indetermination about the dynamical state at the final point $P_2$ is limited to the sphere of radius $p_2$. If we disregard the actual evolution between $t_1$ and $t_2$ but are interested only in the final dynamical state, the motion between point $P_1$ and the generic point $P_2$ can be fictitiously reduced to the green path of figure 1. The path mimics the effect of a Lorentz boost combined with a rotation and the two extremes could be linked by means of a Möbius transformation [20, 21]. This means that, given the correspondence between Möbius transformation and Lorentz transformations, the two momenta $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ at the respective locations $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ could be equally regarded as those momenta of the same event measured by two inertial observers. Note that, in the fictitious evolution or in the Lorentz transformation from $\boldsymbol{p}_1$ to $\boldsymbol{p}_2$, a circular motion appears. It traces an arc over the spherical surface and its stereographic projection is conformal. That is, the stereographic image of a circle (or portion of it) on the sphere is a circle (or an arc) in the equatorial plane. In this way, the 3D problem is scaled down to a simpler 2D problem whose solution is the key to pinpoint $P_2$ on the sphere.

### *3.2. Stereographic projection of the momentum sphere*

The general principles of the stereographic representation are available in the supplementary material and we are going to apply them to the apparent motion of the momentum vector in view of the derivation of the classical Dirac equation.

Although the stereographic technique is nowadays of broad interest [31], its application to solve dynamical problems dates back to the nineteenth century. It forms the premise of the so-called Cayley–Klein approach to the rotation of rigid bodies [18, 19] and has an extension to the treatment of quantum angular momenta [32]. More importantly for our argument, the projection generates the projective representation $SU(2)$ of the rotation group $SO(3)$ [18, 19]. The $SU(2)$ representation is useful for the description of non-relativistic spins [33] and provides the fundamental model to consider for the correspondence between the $SL(2, C)$ group and the Lorentz group used in the QFT approach to the Dirac equation [4, 5]. In this regard, we recall that there is a nice correspondence between the transformation law of the stereographic representation of vectors and the Lorentz transformations (see supplementary material or chapter 1 in [21]).

Crucial aspects of the projection applied to the momentum can be summarized as follows.

The generic point $P$ on the spherical surface defines the generalized momentum $\boldsymbol{p}$ whose absolute value $p$ is established in equation (8). Cartesian coordinates $p_x$, $p_y$ and $p_z$ are visualized with respect to the lab reference frame that defines the North (N) and South (S) poles of the sphere. It is understood that any motion or transformation of the momentum that implies a rescaling of its absolute value $p$, induces a change in the radius of the sphere. This change is taken care in equation (8) and therefore we are only interested in circular motion on the sphere. Thus, we can take advantage of the fact that the projection of a generic circle on the sphere is again a circle in the equatorial plane [20]. This is shown in figure 2 where the arc $C$ that incorporates the generic point $P$ is sent to the projected circle $C_{\text{st}}$ of the equatorial plane that incorporates the projected point $Q$. Its coordinates are $\xi$ and $\eta$. They can be calculated with the help of the sides of similar triangles and are $\xi = p\, p_x/(p - p_z)$ and $\eta = p\, p_y/(p - p_z)$ (see supplementary material). When the correspondence with the complex variable $\zeta = \xi + \mathrm{i}\eta$ (the $y$-axis of figure 2 becomes imaginary) is made and the stereographic matrix of the momentum is introduced

$$M_{\boldsymbol{p}} = \begin{pmatrix} p_z & p_x - \mathrm{i}p_y \\ p_x + \mathrm{i}p_y & -p_z \end{pmatrix}, \tag{9}$$

the structure of the Pauli matrices is immediately recognisable. Indeed, $M_{\boldsymbol{p}}$ is equal to the dot product between the Pauli vector $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $\boldsymbol{p}$. In brief, $M_{\boldsymbol{p}} = \boldsymbol{\sigma} \cdot \boldsymbol{p}$. This matrix is the classical analogue of the helicity operator [2, 5] and, in this respect, has a similar role in the classical context of the current argument.

Here, comes the key point. Whatever change affects the momentum ($\boldsymbol{p} \to \boldsymbol{p}'$), the representation of equation (9) remains unaffected. That is, the Pauli representation of $M_{\boldsymbol{p}}$ is invariant ($M_{\boldsymbol{p}} = \boldsymbol{\sigma} \cdot \boldsymbol{p} \to M_{\boldsymbol{p}'} = \boldsymbol{\sigma} \cdot \boldsymbol{p}'$) and the momentum sphere can always be defined as the determinant of $M_{\boldsymbol{p}}$

$$\det(M_{\boldsymbol{p}}) = -p^2 \to \det(M_{\boldsymbol{p}'}) = -p'^2\,. \tag{10}$$

Next, we see how these peculiar properties can be used to determine the classical Dirac equation.

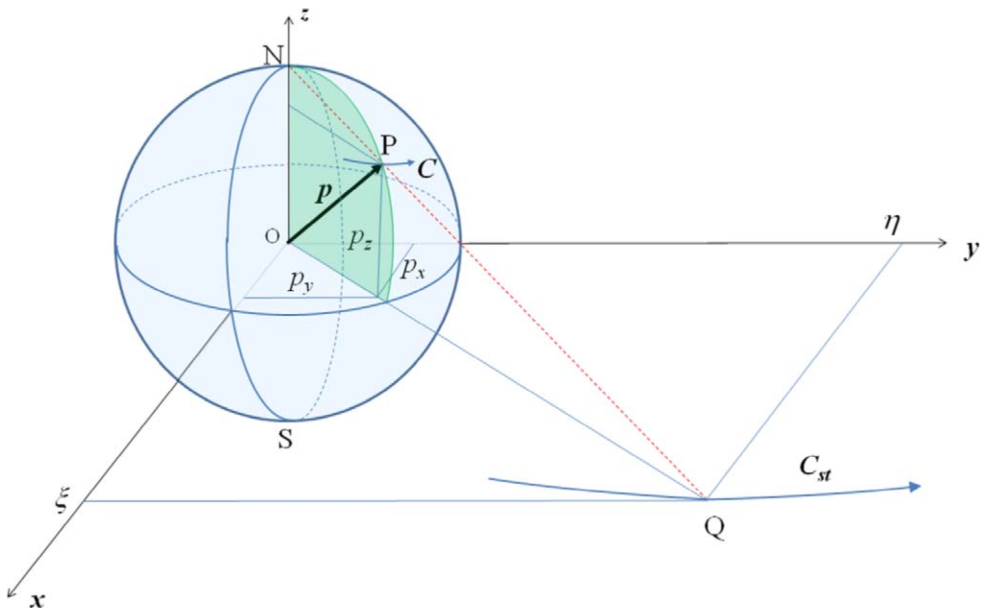


**Figure 2.** Stereographic projection of point $P$ taken on the sphere. The projection is found at the intersection $Q$ between the elongation of the line NP (red line) and the equatorial plane ($x$, $y$). The arc $C$ going through $P$ on the sphere is mapped on the equatorial plane as an equivalent arc $C_{\mathrm{st}}$ of a circle (stereographic image of $C$).

## 4. Classical Dirac equation

The classical description of the momentum sphere has made it possible to link equation (8) with the stereographic matrix of equation (9). But, the question is why do we need such a connection in the first place? It was mentioned before that the stereographic projection takes a fundamental role in the geometric rendering of the Lorentz invariance (see supplementary material or chapter 1 of [21]). The consequence is that the stereographic matrix of equation (9) obeys the principles of relativity under the most general Lorentz transformation that combines boost and rotation [34, 35]. The boost is supposed to change the absolute value of the momentum and the overall effect is already included in equation (8). The stereographic representation takes care of the rotation (e.g. the arc $C$ in figure 2). Note also that the singularities of the projection (North and South poles of figure 2) will not affect the following demonstration that relies on the stereographic matrix. This one becomes diagonal at the singularities without consequences for the demonstration that, in such a circumstance, would include only the $z$ component of the momentum.

### 4.1. Stereographic reduction of the cardinality

Having justified the use of the projection, we can focus on the use of the stereographic matrix in reducing the cardinality (the dimension of the vector space) with the accompanying relocation of the dynamical information. If we start from momentum space with $\dim_{\mathbb{R}}(\mathbb{R}^3) = 3$, we will end up with the spinor space with $\dim_{\mathbb{C}}(\mathbb{C}^2) = 2$ and where the relevant information will be stored.

The eigenvalue problem for the stereographic matrix of the momentum is

$$M_p\begin{pmatrix} u \\ v \end{pmatrix} = \lambda_M \begin{pmatrix} u \\ v \end{pmatrix}. \tag{11}$$

The secular equation yields $\lambda_M = \pm p$. The minus sign means that, in the representation of figure 2, the vector $\boldsymbol{p}$ is in reverse and goes backwards from the spherical surface towards the origin O. The changing sign reveals the indetermination about the direction of the motion and is the manifestation of the uncertainty in the treatment of the motion according to the invariance of the relativistic interval for two different inertial observers. In other words, picking up one specific eigenvalue (or direction in figure 2), its opposite is also plausible.

Depending on the sign, the eigenvectors $\psi_+$ and $\psi_-$ have components $u_\pm$ and $v_\pm$.

$$\psi_\pm = \begin{pmatrix} u_\pm \\ v_\pm \end{pmatrix} \tag{12}$$

and these can be used to demonstrate that their knowledge is sufficient to determine the dynamical state relative to the momentum vector. In other words, the transformation of the 3D problem of equation (8) in a 2D problem of equation (11) induces a one-to-one correspondence between the 3D momentum vector $\boldsymbol{p}$ and the two-component eigenvectors (spinors) of equation (11). For instance, non-trivial solutions to the eigenvalue problem of equation (11) show that one component is made of a complex momentum rotating in the equatorial plane (either $p_x + \mathrm{i}p_y$ or $p_x - \mathrm{i}p_y$). The other component is along the $z$-axis and equals the sum or the difference between $p$ and $p_z$. In other words, $u_\pm$ and $v_\pm$ accommodate complex representations of the original momentum vector $\boldsymbol{p}$. An example is in figure 3 where the generic momentum $\boldsymbol{p}$ of figure 2 has been reproduced together with the eigenvectors $\psi_+$ and $\psi_-$ with components $v_\pm = p_x + \mathrm{i}p_y$ (degenerate) and $u_\pm = p_z \pm p$. It is easy to see that we recover the full $\boldsymbol{p}$ vector if we take $(\psi_+ + \psi_-)/2$. In the end, the decomposition of $\boldsymbol{p}$ in terms of eigenvectors of the stereographic matrix is completely informative of the dynamical state. Of course, other choices for the eigenvectors are possible but they all come down to a linear combination that reproduces the original momentum.

### 4.2. Application of the stereographic reduction

The decomposition of $\boldsymbol{p}$ via its components appearing in the eigenvectors of the stereographic matrix is very useful for our purposes. Indeed, the reduction in the dimensionality of the vector space ensures the solution of the dynamical problem at the cost of encoding the dynamical information in the spinors instead of the original 3D vector. This approach is now applied to the determination of the classical Dirac equation. To show this, we make use of one of the fundamental properties of the Pauli matrices. It involves the $2 \times 2$ identity matrix $I$. That is, $\boldsymbol{p} \cdot \boldsymbol{p} I = (\boldsymbol{\sigma} \cdot \boldsymbol{p})^{\mathbf{2}}$. The right-hand side contains the stereographic matrix and rearranging equation (8) accordingly gives

$$c^2(\boldsymbol{\sigma} \cdot \boldsymbol{p})^2 = (\epsilon - mc^2)(\epsilon + mc^2)I, \tag{13}$$

where $\epsilon = E - V$ has been introduced to simplify the notation. The result is now a matrix equation that can be completed with the multiplication by a vector $\phi$ of components $\phi_1$ and $\phi_2$

$$c^2(\boldsymbol{\sigma} \cdot \boldsymbol{p})^2\phi = (\epsilon - mc^2)(\epsilon + mc^2)\phi. \tag{14}$$

Now, we define a new vector $\chi$ (of components $\chi_1$ and $\chi_2$) according to

$$c(\boldsymbol{\sigma} \cdot \boldsymbol{p})\chi = (\epsilon - mc^2)\phi. \tag{15}$$

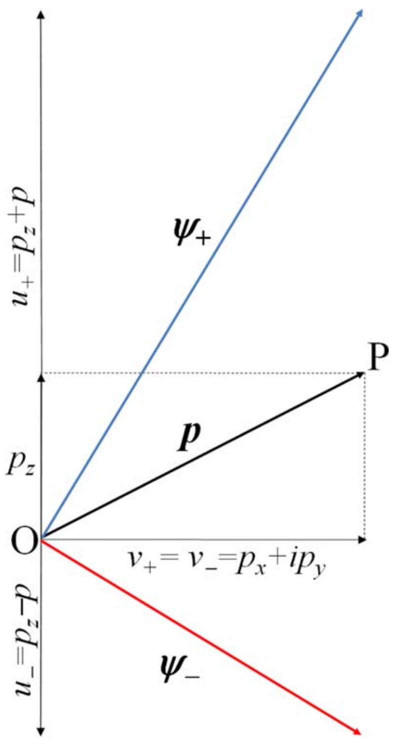


**Figure 3.** Decomposition of the momentum $\boldsymbol{p}$ (black arrow) seen in the plane incorporating the points O, N and P of figure 2. The chosen eigenvectors of the stereographic matrix are $\psi_+$ (blue arrow) and $\psi_-$ (red arrow) with degenerate horizontal component $v_\pm = p_x + \mathrm{i}p_y$ and vertical component $u_\pm = p_z \pm p$.

In this way, equations (14) and (15) define the coupled equations

$$\begin{cases}(\epsilon - mc^2)\phi = c(\boldsymbol{\sigma}\cdot\boldsymbol{p})\chi \\ (\epsilon + mc^2)\chi = c(\boldsymbol{\sigma}\cdot\boldsymbol{p})\phi\end{cases}, \tag{16}$$

where the right-hand sides can be viewed as linear combinations of the eigenvectors of the stereographic matrix $M_{\boldsymbol{p}} = \boldsymbol{\sigma}\cdot\boldsymbol{p}$. The linear system of equation (16) is already recognisable as one manner of writing the Dirac equation (e.g. equation (5) on page 211 in Greiner [2] or equations (8.114) in Sakurai and Napolitano [33]) with $\phi$ and $\chi$ being respectively the so-called large and small components of the bispinor

$$\Psi = \begin{pmatrix}\phi \\ \chi\end{pmatrix}. \tag{17}$$

The use of the bispinor $\Psi$ together with the definition of the matrices $\boldsymbol{\alpha}$ and $\beta$ of equation (4) leads to the classical Dirac equation (see supplementary material)

$$\epsilon\Psi = c(\boldsymbol{\alpha}\cdot\boldsymbol{p})\Psi + mc^2\beta\Psi, \tag{18}$$

where the right side emulates the well-known Dirac Hamiltonian $H_\mathrm{D} = c(\boldsymbol{\alpha}\cdot\boldsymbol{p}) + mc^2\beta$ of equation (3). Understandably, full adherence to that Hamiltonian requires the replacement of the classical momentum with its quantum counterpart.

The procedure described above takes advantage of the spinor fields $\phi$ and $\chi$ that linearize the Hamiltonian. This parallels, of course, the conventional RQM and QFT treatments of the Dirac equation. However, the conventional QFT approach is built around the Lorentz transformations for the spinor fields and this strategy leads to the correct equation [3–5]. Here and similarly in the RQM approach [1, 2, 6], the procedure disregards the transformation properties of the bispinor and gains from the Lorentz invariance of the stereographic representation [21]. This is enough to get a covariant equation with transformation properties of the bispinor adjusted to those of the stereographic representation.

## 5. Comparison with conventional approaches

In the derivation of equations (16) or (18) we spared ourselves any quantum switch. Now, we are all set to turn the classical result into the well-known Dirac equation. This is accomplished thanks to the transformation of the actual momentum vector in its quantum-mechanical representation $\boldsymbol{p}_0 = -\mathrm{i}\hbar\boldsymbol{\nabla}$ (generally different from the canonical momentum) together with the analogous representation of the Hamiltonian $H = E = \mathrm{i}\hbar\partial/\partial t$. These replacements yield

$$(\mathrm{i}\hbar\partial/\partial t - V)\Psi = (c\boldsymbol{\alpha}\cdot\boldsymbol{p} + mc^2\beta)\Psi, \tag{19}$$

where $\boldsymbol{\alpha}$ and $\beta$ are given in equation (4) and $\boldsymbol{p} = -\mathrm{i}\hbar\boldsymbol{\nabla} + \boldsymbol{P}$ incorporates the correction $\boldsymbol{P}$ for the vector potential. To recover the common Dirac equation found in the appropriate literature, the substitutions for free fields ($V = 0$ and $\boldsymbol{p} = \boldsymbol{p}_0$ or $\boldsymbol{P} = 0$) are in order.

Some comments are fitting at this point. In the summary of section 2, it was remarked that RQM and QFT procedures are inevitably set in the free space. External fields disrupt the homogeneity of each point in space (translational symmetry) and this is detrimental to the quantum-mechanical argument. By contrast, the stereographic approach does not need the homogeneity of the free space so essential in conventional treatments. The same goes for the relativistic linearity of the Dirac equation. If the RQM and QFT derivations take the linearity as an external constraint that dictates the correct structure of the equation, the stereographic approach works in reverse. The linearity emerges as an internal feature of the projective representation that shifts the dynamical information from the original 3D momentum space to the 2D spinor space. In the end, none of the hypotheses made in RQM and QFT procedures are replicated here.

Remarkably, the stereographic approach shows that the spin matrices do not promote the picture of a particle spinning like a top. The denial agrees with the answer to the question ‘What is spin?’ reported by Ohaian [36] who observes a classical analogy with the angular momentum carried by a circularly polarized wave. Similarly, a classical analogue of the spin is here visualized in the stereographic representation of the apparent rotation that the most general Lorentz transformations introduce in the momentum between two inertial reference frames. This fictitious rotation is visually reproduced in the arc of figure 2 where the conformal stereographic map sends 3D rotational paths to 2D rotations in the equatorial plane. In other words, being the rotation not real, the apparent motion of the momentum does not correspond to a physical reality of a spinning particle and is solely caused by the representation scheme based on the Pauli matrices (stereographic spin). On the other hand, if the interaction entails a true particle rotation occurring in real space, there are components of the generalized momentum coming from the orbital motion. Then, the real orbital momentum couples with the stereographic spin through the generalized momentum and nevertheless, the two momenta keep their distinction. The spin is again an effect of the projective representation whereas the orbital momentum is the manifestation of a true rotation. It is their coupling in $\boldsymbol{\alpha} \cdot \boldsymbol{p}$ of equation (18) that creates the common idea of an 'intrinsic' angular momentum capable of physical effects (a couple of examples are in the supplementary material).

## 6. Conclusions

In conclusion, the Dirac equation is found after a simple geometric procedure that is independent from those assumptions made in conventional contexts. The method is based on the stereographic projection that is known to produce the Pauli matrices in the Cayley–Klein description of rotating rigid bodies of classical mechanics. Its application to the generalized momentum solves the dynamical problem posed by the relativistic energy-momentum relationship where the absolute value of the momentum is known but its direction is indeterminate. However, the indeterminacy is constrained to a sphere whose stereographic projection has a conformal image in the complex plane coincident with the equatorial plane. The associated eigenvalue problem set for the stereographic matrix reduces the 3D physics to a simpler 2D context. The operation results in the classical Dirac equation. Thanks to this preparatory material, the transition to the original quantum equation becomes effortless in comparison to conventional approaches.

The procedure illustrated in this work can be understood by students who are curious about the Dirac equation but have modest knowledge of quantum mechanics. For students acquainted with RQM and QFT arguments, the geometric viewpoint of the stereographic approach can complement their educational background coming from the tools of abstract algebra and elaborated analysis that are typical of QFT and RQM routes to the equation.

## Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

## ORCID iDs

Michele Marrocco https://orcid.org/0000-0002-5279-9497

## References

[1] Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)
[2] Greiner W 2000 *Relativistic Quantum Mechanics* (Berlin: Springer)
[3] Itzykson C and Zuber J-B 1980 *Quantum Field Theory* (New York: McGraw-Hill)
[4] Weinberg S 1995 *The Quantum Theory of Fields* (Cambridge: Cambridge University Press)
[5] Peskin M E and Schroeder D V 1995 *An Introduction to Quantum Field Theory* (London: CRC Press)
[6] Dirac P A M 1928 The quantum theory of the electron *Proc. Roy. Soc.*A **117** 610–24
[7] Lanczos C 1929 Die tensoranalytischen beziehungen der diracschen gleichung *Zeits. f. Phys.* **57** 447–73
[8] Conway A W 1937 Quanternion treatment of the relativistic wave equation *Proc. R. Soc.*A **162** 145–54
[9] Srinivasanan S K and Sudarshan E C G 1996 A direct derivation of the Dirac equation via quaternion measures *J. Phys. A: Math. Gen.* **29** 5181–6

[10] Célérier M-N and Nottale L 2004 Quantum–classical transition in scale relativity *J. Phys. A: Math. Gen.* **37** 931–55
[11] Feynman R P and Hibbs A R 2010 *Quantum Mechanics and Path Integrals* (New York: Dover)
[12] Bess L 1979 A diffusion model for the Dirac equation *Found. Phys.* **9** 27–54
[13] Cufaro Petroni N and Vigier J P 1981 Stochastic derivation of the Dirac equation in terms of a fluid spinning tops endowed with random fluctuations at the velocity of light *Phys. Lett. A* **81** 12–4
[14] Gaveau B, Jacobson T, Kac M and Schulman L S 1984 Relativistic extension of the analogy between quantum mechanics and Brownian motion *Phys. Rev. Lett.* **53** 419–22
[15] Kudo T, Ohba I and Nitta H 2001 A derivation of the Dirac equation in an external field based on the Poisson process *Phys. Lett.* A **286** 227–30
[16] Jack N Y and van Dam H 2003 A geometrical derivation of the Dirac equation *Phys. Lett. A* **309** 335–9
[17] Santamato E and De Martini F 2013 Derivation of the Dirac equation by conformal differential geometry *Found. Phys.* **43** 631–41
[18] Goldstein H 1980 *Classical Mechanics* (Reading: Addison-Wesley)
[19] José J V and Saletan E J 2013 *Classical Dynamics: A Contemporary Approach* (Cambridge: Cambridge University Press)
[20] Needham T 2012 *Visual Complex Analysis* (Oxford: Oxford University Press)
[21] Penrose R and Rindler W 1986 Spinors and space-time *Two-Spinor Calculus and Relativistic Field* vol 1 (Cambridge: Cambridge University Press)
[22] Villalba V M 1994 The angular momentum operator in the Dirac equation *Eur. J. Phys.* **15** 191–6
[23] Gaioli F H and Garcia Alvarez E T 1995 Some remarks about intrinsic parity in Ryder's derivation of the Dirac equation *Am. J. Phys.* **63** 177
[24] Lerner L 1996 Derivation of the Dirac equation from a relativistic representation of spin *Eur. J. Phys.* **17** 172–5
[25] Holstein B R 1997 Diagonalization of the Dirac equation: an alternative approach *Am. J. Phys.* **65** 519–22
[26] Alberto P, Fiolhais M and Oliveira M 1998 On the relativistic L–S coupling *Eur. J. Phys.* **19** 553–62
[27] Nikolić H 2014 How (not) to teach Lorentz covariance of the Dirac equation *Eur. J. Phys.* **35** 035003
[28] Vaz J Jr 2016 The Clifford algebra of physical space and Dirac theory *Eur. J. Phys.* **37** 055407
[29] Alberto P, Fiolhais C and Gil V M S 1996 Relativistic particle in a box *Eur. J. Phys.* **17** 19–22
[30] Ram B 1982 Quark confining potential in relativistic equations *Am. J. Phys.* **50** 549–51
[31] https://en.wikipedia.org/wiki/Stereographic_projection
[32] Biedenharn L C and Louck J D 1985 Angular momentum in quantum physics *Encyclopedia of Mathematics and Applications* ed G-C Rota vol 8 (Cambridge: Cambridge University Press)
[33] Sakurai J J and Napolitano J 2021 *Modern Quantum Mechanics* (Cambridge: Cambridge University Press)
[34] Ferraro R and Thibeault M 1999 Generic composition of boosts: an elementary derivation of the Wigner rotation *Eur. J. Phys.* **20** 143–51
[35] O'Donnel K and Visser M 2011 Elementary analysis of the special relativistic combination of velocities, Wigner rotation and Thomas precession *Eur. J. Phys.* **32** 1033–47
[36] Ohanian H C 1986 What is spin? *Am. J. Phys.* **54** 500–5