

But what is spin, really?

Louis-Jigmé Petit*
Faculty of Physics and Engineering, University of Strasbourg

Starting from the importance of classical angular momentum, we will undertake a journey motivated by history towards an intuitive understanding of the necessity for an intrinsic angular momentum (or spin) of particles. The aim of this paper is to go through the entire process of the discovery of spin, as if the reader was discovering it himself. To this end, the procedure of quantisation of angular momentum and the main conceptual interests of the operators that arise from it will be outlined, and evidence of the existence of half integer quantum numbers will be presented. The rest of this paper will be focused on deriving and exploring Pauli's theory of spin in depth, while we learn to let go of our classical understanding of the quantum realm in favour of an intuition for the seemingly nonsensical. Finally, consequences of this new intrinsic degree of freedom will be explored by studying the symmetrization postulate and its origins in the spin-statistics theorem, giving rise to the well known properties of Fermions and Bosons.

I. INTRODUCTION

After the discovery of the electron by J.J. Thompson in 1897 and that of the nucleus by Rutherford in 1911, it was clear that the atom did not behave as one irreducible object, but as a collection of electrons orbiting a nucleus, somehow. This "somehow" has been one of the starting points of the entire field of quantum mechanics, as it was clear that simple orbits as those of planets around a star could not explain the atom. One of the reasons why is that moving charges radiate and lose energy, which would cause the radius of the orbit to become smaller over time, but electrons seem to stubbornly insist not to crash into the nucleus. Another reason was that atomic spectra indicated discrete allowed energy levels for electrons, even though classical orbits can produce a continuous range of energies.

But the orbit idea could not be so easily abandoned, which is why during the 1910s Bohr developed his famous model of the atom by quantizing the values of perhaps the most useful physical quantity when it comes to the study of orbits: angular momentum.

Even though his model turned out to be wrong on several accounts, the central role of angular momentum in the study of the atom, and later quantum systems in general, was there to stay.

Indeed, in the 1920's the modern framework of quantum mechanics was being developed by Schrödinger, Heisenberg and many others, and the problem of the Hydrogen atom was solved, or so it seemed. Despite the uncanny predictive power of the new model, especially for the time, a few phenomena seemed to escape its grasp. Chief among those was the fact that the electron seemed to have twice as many states as the theory predicted, and that somehow no more than a single electron was allowed to occupy each of these states. Weirder still, the degree of freedom of these new states seemed to behave precisely as an angular momentum with very peculiar properties thought to be no more than mathematical artefacts until then. These were the first signs of a new fundamental degree of freedom that had never before been observed in human history: quantum spin.

In order to understand the behaviour of this "intrinsic angular momentum", we will start by thoroughly going through the procedure of quantisation of angular momentum, before stating the historical arguments in favour of the necessity for this new physical quantity. This will allow us to study and understand Pauli's empirical theory of spin and to explore its consequences. We will also briefly go beyond the empirical theory and give a qualitative explanation of the necessity of spin in Dirac's theory of relativistic quantum mechanics and the generalisation of Pauli's exclusion principle with the spin statistics theorem from quantum field theory.

II. FROM CLASSICAL TO QUANTUM ANGULAR MOMENTUM

In classical mechanics, the angular momentum $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ is a quantity that can be seen from two complementary standpoints:

- A dynamical property constructed from $\boldsymbol{r}$ and $\boldsymbol{p}$, analogous to linear momentum for circular paths, that becomes extremely useful when dealing with rotations in the motion of a system

* louis-jigme.petit@etu.unistra.fr

- A fundamental quantity that is conserved entirely for a closed system as a consequence of the isotropy of space, and conserved in some capacity for a system exhibiting some continuous rotational symmetry

This second interpretation can be shown quite easily by considering that the Hamiltonian of a rotationally invariant system should remain the same after a rotation.

Let $\boldsymbol{n}$ be a unit vector along an axis with respect to which the considered system is rotationally invariant, and $\delta\boldsymbol{\varphi} = \delta\varphi\,\boldsymbol{n}$ be an infinitesimal rotation of angle $\delta\varphi$ around that axis. The rotated vectors $\boldsymbol{r'}$ and $\boldsymbol{p'}$ can be expressed as functions of their unrotated counterparts $\boldsymbol{r}$ and $\boldsymbol{p}$ as:

$$\boldsymbol{r'} = \boldsymbol{r} + \delta\boldsymbol{r} = \boldsymbol{r} + \delta\boldsymbol{\varphi} \times \boldsymbol{r} \qquad\qquad \boldsymbol{p'} = \boldsymbol{p} + \delta\boldsymbol{p} = \boldsymbol{p} + \delta\boldsymbol{\varphi} \times \boldsymbol{p}$$

Therefore, assuming the Hamiltonian of the system remains unchanged, one can write:

$$\begin{aligned} H'(\boldsymbol{r'},\boldsymbol{p'}) = H(\boldsymbol{r},\boldsymbol{p}) &\implies H(\boldsymbol{r}+\delta\boldsymbol{r},\boldsymbol{p}+\delta\boldsymbol{p}) = H(\boldsymbol{r},\boldsymbol{p}) \\ &\implies H(\boldsymbol{r},\boldsymbol{p}) + \boldsymbol{\nabla_r} H(\boldsymbol{r},\boldsymbol{p})\,\delta\boldsymbol{r} + \boldsymbol{\nabla_p} H(\boldsymbol{r},\boldsymbol{p})\,\delta\boldsymbol{p} = H(\boldsymbol{r},\boldsymbol{p}) \end{aligned} \tag{1}$$

Using Hamilton's equation, i.e. $\dfrac{\partial H}{\partial q_i} = -\dot{p}_i$, $\dfrac{\partial H}{\partial p_i} = \dot{q}_i$, (1) becomes:

$$\begin{aligned} (1) \implies -\dot{\boldsymbol{p}} \cdot \delta\boldsymbol{r} + \dot{\boldsymbol{r}} \cdot \delta\boldsymbol{p} = 0 &\implies -\dot{\boldsymbol{p}} \cdot \delta\boldsymbol{\varphi} \times \boldsymbol{r} + \dot{\boldsymbol{r}} \cdot \delta\boldsymbol{\varphi} \times \boldsymbol{p} = 0 \\ &\implies \dot{\boldsymbol{p}} \times \boldsymbol{r} \cdot \delta\boldsymbol{\varphi} - \dot{\boldsymbol{r}} \times \boldsymbol{p} \cdot \delta\boldsymbol{\varphi} = 0 \\ &\implies (\boldsymbol{r} \times \dot{\boldsymbol{p}} + \dot{\boldsymbol{r}} \times \boldsymbol{p}) \cdot \delta\boldsymbol{\varphi} = 0 \\ &\implies \delta\varphi\,\boldsymbol{n} \cdot \frac{d}{dt}(\boldsymbol{r} \times \boldsymbol{p}) = 0 \qquad\qquad \forall\,\delta\varphi \\ &\implies \frac{d}{dt}(\boldsymbol{n} \cdot \boldsymbol{r} \times \boldsymbol{p}) \equiv \frac{d}{dt}(\boldsymbol{n} \cdot \boldsymbol{L}) = 0 \end{aligned}$$

Therefore, $\boldsymbol{n} \cdot \boldsymbol{L}$ is a constant of motion if $\boldsymbol{n}$ is along an axis of symmetry of the system[1]. As one example among many of the usefulness of this quantity, the conservation of angular momentum is precisely why orbits happen on a plane, and it is the mathematical way of stating Kepler's law of areas!

Seeing how useful this quantity is in the classical framework, one can wonder if an analogous object exists in the quantum theory, and if so if it is as useful as in the classical case.

As we will see, the answer to both of these questions is affirmative. Actually, one could argue that it is even more important in the quantum framework than in the classical one. But let's not get ahead of ourselves, we still need to find out what this object is in order to study its behaviour.

In quantum mechanics, physical quantities are replaced with hermitian operators, called observables, acting on a Hilbert space associated with the system that is being studied, often referred to as the state space. Each element of this space contains all the information related to the corresponding physical state (which is why we will usually simply refer to these elements as "states"). To each observable, we can associate a basis of eigenstates (or eigenbasis), which generates the entire associated Hilbert space.

There isn't always a simple approach to finding the observable corresponding to a classical dynamical variable, although it can often be found by applying the canonical quantisation procedure.

$$f(\boldsymbol{r},\boldsymbol{p},t) \to F(\boldsymbol{R},\boldsymbol{P},t)$$

Where $\boldsymbol{R}$ and $\boldsymbol{P}$ are respectively the position and momentum operators and F is a hermitian operator written as a function of the same form as f with non commuting products properly symmetrised (for example $x\,p_x \to (X\,P_x + P_x\,X)/2)$[2].

Applying this procedure to angular momentum gives $\boldsymbol{L} = \boldsymbol{R} \times \boldsymbol{P}$ [3] (where all the products are of the form R_iP_j, $i \neq j$, which commute so no symmetrisation is necessary).

---

[1] The extension of the proof to a system of particles presents no conceptual difficulty and is left to the reader, who should find that what is conserved is in fact the sum of all angular momenta of particles.

[2] Another requirement of this procedure is the transformation of Poisson brackets into commutators, which is part of the reason why it is not always successful as it imposes heavy constraints of the linear map F, but this is beyond the scope of this paper.

[3] From now on, the symbol $\boldsymbol{L}$ and the associated L_i will stand for the quantum operator of angular momentum, and not the corresponding classical dynamical variable.

Although this procedure gives the right formula, it completely hides the reason why this is true, and gives no insight on which properties are transferred from classical mechanics or not. Since the object we started from had a lot to do with rotations and rotational symmetry, one can reasonably wonder if we can find this operator by considering the conservation of a rotationally symmetric Hamiltonian under an infinitesimal rotation.

Consider a wavefunction $\psi(\boldsymbol{r})$ in a given coordinate system. A rotation of the wavefunction of an angle $\delta\varphi$ around the axis spanned by $\boldsymbol{n}$ is equivalent to the rotation of the coordinate system by an angle $-\delta\varphi$ around the same axis (i.e. active transformation vs passive), therefore:

$$\begin{aligned}\mathcal{R}_{\boldsymbol{n}}(\delta\varphi)\ \psi(\boldsymbol{r}) = \psi(\mathcal{R}_{\boldsymbol{n}}(-\delta\varphi)\,\boldsymbol{r}) &= \psi(\boldsymbol{r} - \delta\varphi\,\boldsymbol{n}\times\boldsymbol{r}) \\ &= (1 - \delta\varphi\,\boldsymbol{n}\times\boldsymbol{r}\cdot\boldsymbol{\nabla})\psi(\boldsymbol{r}) \\ &= (1 - \delta\varphi\,\boldsymbol{n}\cdot\boldsymbol{r}\times\boldsymbol{\nabla})\psi(\boldsymbol{r}) \\ &= \left(1 - \frac{i}{\hbar}\delta\varphi\,\boldsymbol{n}\cdot\boldsymbol{r}\times(-i\hbar\boldsymbol{\nabla})\right)\psi(\boldsymbol{r})\end{aligned} \tag{2}$$

Therefore $R_{\boldsymbol{n}}(\delta\varphi) = 1 - \frac{i}{\hbar}\delta\varphi\,\boldsymbol{n}\cdot\boldsymbol{R}\times\boldsymbol{P}$ is the general form of the operator for an infinitesimal rotation about $\boldsymbol{n}$, which should leave a Hamiltonian that is rotationally symmetric around this axis invariant.

$$\begin{aligned}R_{\boldsymbol{n}}(\delta\varphi)^{\dagger} H\, R_{\boldsymbol{n}}(\delta\varphi) = H \implies& \left(1 + \frac{i}{\hbar}\delta\varphi\,\boldsymbol{n}\cdot\boldsymbol{R}\times\boldsymbol{P}\right) H \left(1 - \frac{i}{\hbar}\delta\varphi\,\boldsymbol{n}\cdot\boldsymbol{R}\times\boldsymbol{P}\right) = H \\ \implies& H + \frac{i}{\hbar}\delta\varphi\,(\,(\boldsymbol{n}\cdot\boldsymbol{R}\times\boldsymbol{P})\,H - H\,(\boldsymbol{n}\cdot\boldsymbol{R}\times\boldsymbol{P})\,) + O(\delta\varphi^2) = H \\ \implies& [\boldsymbol{n}\cdot\boldsymbol{R}\times\boldsymbol{P}, H] = 0 \\ \implies& \frac{d}{dt}\,\langle\boldsymbol{n}\cdot\boldsymbol{R}\times\boldsymbol{P}\rangle \equiv \frac{d}{dt}\,\langle\boldsymbol{n}\cdot\boldsymbol{L}\rangle = 0\end{aligned}$$

Where the last implication is made possible by Ehrenfest's theorem.

We have not only found that our initial guess was in fact correct, but that rotational invariance still conserves angular momentum along the axes of symmetry (in the quantum sense at least, hence the expectation value). This new operator therefore has strong conceptual links with its classical counterpart, as is to be expected, but it also has its share of differences, which will become very important to our story.

III. PROPERTIES OF THE QUANTUM ANGULAR MOMENTUM

First of all, it is important to notice that $\boldsymbol{L}$ is a Hermitian operator, which makes its eigenvalues real, measurable quantities. To find these eigenvalues, it is necessary to diagonalise $\boldsymbol{L}$.

Unfortunately, we are going to have a problem. Up until now, one could have thought that the angular momentum operator could have behaved similarly to the $\boldsymbol{R}$ and $\boldsymbol{P}$ operators, for which all components could be diagonalised simultaneously, and end up having an "angular momentum representation" of the Hilbert space. But these dreams are crushed by a fact so fundamental that it will end up replacing the definition of angular momentum altogether: the components of $\boldsymbol{L}$ do not commute!

$$\begin{aligned}[L_x, L_y] &= [Y\,P_z - Z\,P_y\,,\ Z\,P_x - X\,P_z] \\ &= [Y\,P_z, Z\,P_x] - [Y\,P_z, X\,P_z] - [Z\,P_y, Z\,P_x] + [Z\,P_y, X\,P_z] \\ &= Y\,P_x[P_z, Z] + P_y\,X[Z, P_z] \\ &= [P_z, Z](Y\,P_x - X\,P_y) \\ &= i\hbar\,L_z\end{aligned}$$

It can be checked easily that this generalises to[1]:

[1] Note to the reader: Einstein summation notation will be used throughout this paper.

$$[L_i, L_j] = i\hbar\,\varepsilon_{ijk}\,L_k \tag{3}$$

This is the most important fact about quantum angular momenta, and it is not all that surprising after taking a step back.

If we go back to the rotation operator and notice that successive rotations around the same axis can be written as $R_{\boldsymbol{n}}(\alpha)R_{\boldsymbol{n}}(\beta) = R_{\boldsymbol{n}}(\alpha+\beta)$, we can write that for a finite rotation $R_{\boldsymbol{n}}(\theta) = R_{\boldsymbol{n}}(\theta/k)^k$ for any integer k. Taking the limit $k \to \infty$ thus allows us to use the formula for an infinitesimal rotation we found earlier.

$$R_{\boldsymbol{n}}(\theta) = \lim_{k\to\infty} R_{\boldsymbol{n}}\left(\frac{\theta}{k}\right)^k = \lim_{k\to\infty}\left(1 - i\frac{\theta}{k}\frac{\boldsymbol{n}\cdot\boldsymbol{L}}{\hbar}\right)^k = e^{-\frac{i}{\hbar}\theta\,\boldsymbol{n}\cdot\boldsymbol{L}} \tag{4}$$

This can be seen as the definition of the operator $\boldsymbol{L}$, which can be referred to as the *generator* of rotations, as it wholly encodes the behaviour of rotations along any axis $\boldsymbol{n}$. In position representation, one has (with slight abuse of notation)$\boldsymbol{n}\cdot\boldsymbol{L} = -i\hbar\,\partial_{\theta_{\boldsymbol{n}}}$, where $\theta_{\boldsymbol{n}}$ is simply a notation for the angle of rotation around the axis given by $\boldsymbol{n}$. It is in this sense that $\boldsymbol{L}$ *generates* rotation, as it gives the variation of the state under an infinitesimal rotation by differentiating it. In group theory language, the $\boldsymbol{n}\cdot\boldsymbol{L}$ are elements of the *algebra* of rotations and the R are elements of the *group*[1].

In practice, this means that $\boldsymbol{n}\cdot\boldsymbol{L}$ can be used to study rotations around the axis $\boldsymbol{n}$, but not around other axes. In other words $\boldsymbol{n}\cdot\boldsymbol{L}$ isn't the "projection of the full angular momentum along $\boldsymbol{n}$" but the full angular momentum operator measured along $\boldsymbol{n}$ [1]. The components of $\boldsymbol{L}$ form a basis of the algebra, i.e. a basis of the space that describes infinitesimal rotations along any axis, and the vector operator is just a neat way to collect this basis in a single object, just as writing $(\hat{x}, \hat{y}, \hat{z})$ would be a neat (but slightly useless) way to collect the basis vectors of real space in a single object (that isn't really a vector).

This can be linked back to $\boldsymbol{R}$ and $\boldsymbol{P}$. Both can be associated with translation (respectively in momentum and real space) in the same way $\boldsymbol{L}$ can be linked to rotations by (4), and one can notice that translations only act on the axis under consideration (i.e. a translation along x leaves y and z invariant), whereas rotations only act on the others (i.e. a rotation along the x axis only leaves x invariant, while mixing up the components of y and z). This ensures that position and momentum operators along orthogonal axes must commute, whereas angular momentum operators must not.

All of this can serve as an intuitive basis for understanding angular momentum, reframing its peculiar quantum behaviour as a simple consequence of the non commutative behaviour of rotations (see Figure 1).

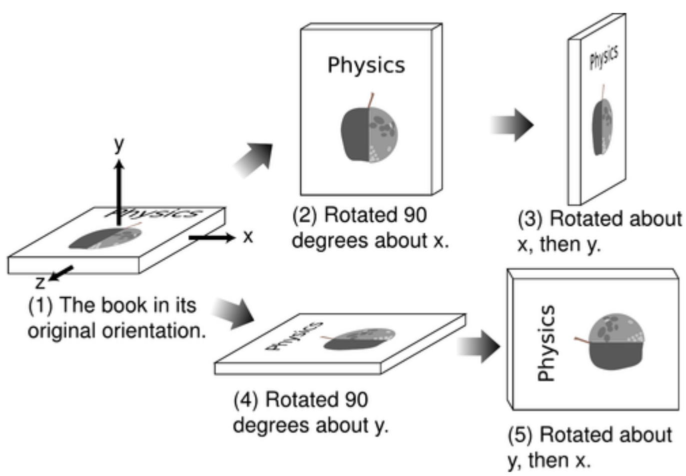

FIG. 1: Non commutativity of rotations, from General Relativity, p. 256. by Benjamin Crowell

[1] In mathematical terms, the algebra is the tangent space of the group at the identity, in layman's terms, the algebra focuses on infinitesimal transformation, which are generally simpler to study and can be mapped back to the finite transformations through the exponential map (4).

All of this is to say that as the different L_i do not commute, they cannot be simultaneously diagonalised, an arbitrary axis is therefore chosen. By convention, we will study L_z, the angular momentum operator along the z axis. We will denote α its eigenvalues and $|\alpha\rangle$ its corresponding eigenvectors.

But now we have a second problem. Space is supposed to be isotropic, so the choice of the axis shouldn't matter in our description of angular momentum. But L_z and, say, L_x will in general not have the same eigenvector decomposition for a given state, as they are non commuting. We can therefore infer that studying angular momentum along one axis is not sufficient to describe it completely.

In fact we would like to have an operator built from the angular momentum operators that is invariant under rotation, and thus commutes with angular momentum about all axes.

By considering an arbitrary polynomial and applying these constraints, one finds that the simplest operator satisfying all of them is:

$$\boldsymbol{L}^2 = L_x^2 + L_y^2 + L_z^2 \tag{5}$$

One can also show that all other operators satisfying the constraints can be written as functions of this one, making this choice natural.

This operator can be seen as measuring the quantum fluctuations of angular momentum in an isotropic way, which in a way quantifies the modulus of the angular momentum[1]. But once again one must be careful with this interpretation, as $\boldsymbol{L}$ is not a simple vector but just a way to write a basis of operators.

Since the L_i commute with $\boldsymbol{L}^2$, it is possible to diagonalise it at the same time as L_z. We can therefore write the eigenvalues of $\boldsymbol{L}^2$ as β and the vectors of the eigenbasis common to it and L_z as $|\beta, \alpha\rangle$.

As the aim is to characterise the behaviour of angular momentum as a whole, it is interesting to check if $L_{x,y}\,|\beta, \alpha\rangle$ are eigenkets of $\boldsymbol{L}^2$ and L_z.

$$\boldsymbol{L}^2(L_{x,y}\,|\beta, \alpha\rangle) = L_{x,y}\boldsymbol{L}^2\,|\beta, \alpha\rangle = \beta\, L_{x,y}\,|\beta, \alpha\rangle$$

$$L_z(L_x\,|\beta, \alpha\rangle) = ([L_z, L_x] + L_x L_z)\,|\beta, \alpha\rangle = (i\hbar L_y + \alpha L_x)\,|\beta, \alpha\rangle$$

$$L_z(L_y\,|\beta, \alpha\rangle) = ([L_z, L_y] + L_y L_z)\,|\beta, \alpha\rangle = (-i\hbar L_x + \alpha L_y)\,|\beta, \alpha\rangle$$

As expected because of the commutation properties of the operators, $L_{x,y}\,|\beta, \alpha\rangle$ are eigenkets of $\boldsymbol{L}^2$ but not of L_z. But checking this was not entirely pointless, as it can be noticed that the operators defined as the linear combinations $L_\pm = L_x \pm iL_y$ have some very interesting properties with respect to L_z!

$$\begin{aligned} L_z(L_\pm\,|\beta, \alpha\rangle) = (L_z L_x \pm i L_z L_y)\,|\beta, \alpha\rangle &= (i\hbar L_y + \alpha L_x \pm \hbar L_x \pm i\alpha L_y)\,|\beta, \alpha\rangle \\ &= (\alpha \pm \hbar)(L_x \pm i L_y)\,|\beta, \alpha\rangle \\ &= (\alpha \pm \hbar) L_\pm\,|\beta, \alpha\rangle \end{aligned} \tag{6}$$

$L \pm |\beta, \alpha\rangle$ is therefore an eigenket of L_z of eigenvalue $\alpha \pm \hbar$, meaning that we can write:

$$L_\pm\,|\beta, \alpha\rangle = C_\pm(\alpha, \beta)\,|\beta, \alpha \pm \hbar\rangle \tag{7}$$

Where $C_\pm(\alpha, \beta)$ are constants to be determined. $L_\pm$ are called ladder operators because they raise or lower the eigenvalue of L_z by $\hbar$. Furthermore, these operators are not Hermitian, but Hermitian conjugates of one another, i.e. $L_+^\dagger = L_-$ and the other way around.

The fact that operators with such properties should exist is not at all trivial[2], it is a very lucky twist of fate that they appear naturally in our derivation of the physical properties of angular momentum, and they will greatly simplify (and simply make possible) our study of the eigenbasis.

By noticing that we can write $\boldsymbol{L}^2$ and L_z as functions of $L_\pm$, we will have all the tools in hand to properly study this eigenbasis.

---

[1] One could ask if $\boldsymbol{L}^2$ alone could be enough to study angular momentum, but it is important to remember that this operator only measures the isotropic fluctuations of angular momentum, but what we want is to measure it along an axis, hence the importance of keeping L_z.

[2] For the interested reader, in the group theory that underlies the study of angular momenta, they come from the complexification of the Lie algebra $\mathfrak{su}(2)$, which we are secretly studying from a physicists point of view.

$$L_\pm L_\mp = (L_x \pm iL_y)(L_x \mp iL_y) = L_x^2 + L_y^2 \pm i[L_y, L_x] = L_x^2 + L_y^2 \pm \hbar L_z$$

$$\implies \begin{cases} \boldsymbol{L}^2 = L_z^2 + \dfrac{1}{2}(L_+L_- + L_-L_+) \\ L_z = \dfrac{1}{2\hbar}(L_+L_- - L_-L_+) \end{cases} \tag{8}$$

By applying (8) to the eigenvalue equations $\boldsymbol{L}^2 |\beta, \alpha\rangle = \beta |\beta, \alpha\rangle$ and $L_z |\beta, \alpha\rangle = \alpha |\beta, \alpha\rangle$ and using (7), one finds the general relation relating the eigenvalues α and β to the constants $C_\pm(\alpha, \beta)$:

$$C_\pm(\alpha, \beta)\, C_\mp(\alpha \pm \hbar, \beta) = \beta - \alpha(\alpha \pm \hbar) \tag{9}$$

Using the fact that $|L_\pm |\beta, \alpha\rangle|^2 = C_\pm(\alpha, \beta)^2$ and (9), gives:

$$|L_\pm |\beta, \alpha\rangle|^2 = \langle \beta, \alpha| L_\mp L_\pm |\beta, \alpha\rangle = C_\pm(\alpha, \beta)\, C_\mp(\alpha \pm \hbar, \beta) = C_\pm(\alpha, \beta)^2$$

$$\implies \beta - \alpha(\alpha \pm \hbar) = C_\pm(\alpha, \beta)^2 \tag{10}$$

The information given by (10) is twofold, first we have found an expression for the constants $C_\pm(\alpha, \beta) = \sqrt{\beta - \alpha(\alpha \pm \hbar)}$, and second and most important, the fact that the modulus square must be positive enforces:

$$\beta - \alpha(\alpha \pm \hbar) \geqslant 0 \implies \alpha(\alpha \pm \hbar) \leqslant \beta \tag{11}$$

Which means there is an upper and lower bound to α[1] for which $C_+(\alpha_{max}(\beta), \beta) = 0$ and $C_-(\alpha_{min}(\beta), \beta) = 0$.

$$\alpha_{max} = \frac{-\hbar + \sqrt{\hbar^2 + 4\beta}}{2} = -\alpha_{min} \tag{12}$$

Putting everything together now, we know that there exists an upper and lower bound for α, but since the ladder operators raise the eigenvalues α by $\hbar$ each time they are applied and only annihilate states with precisely these higher and lower bounds for α, there must be an integer multiple of $\hbar$ between these bounds in order to stay between them by repeated action of $L_\pm$ (see Figure 2).

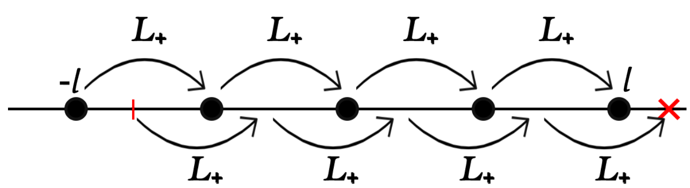

FIG. 2: The eigenvalues of L_z must lie on equally spaced points between l and $-l$, else $L_\pm$ will shift the eigenvalues out of the allowed bounds

This also means that α can only take discrete values separated by $\hbar$, since if this was not the case applying L_+ would eventually give a value of $\alpha > \alpha_{max}$. Using (12), this can be written as:

$$\alpha_{max} = \alpha_{min} + k\hbar = -\alpha_{max} + k\hbar \implies \alpha_{max} = \frac{k}{2}\hbar \qquad k \in \mathbb{N} \tag{13}$$

[1] Notice that for $\beta = 0$, these two bounds must meet and $\alpha = 0$ since both conditions $\alpha(\alpha + \hbar) \leq 0$ and $\alpha(\alpha - \hbar) \leq 0$ must be met independently, thus excluding the values $\alpha = \pm\hbar$.

Plugging (13) into (12), we finally get:

$$\frac{-\hbar+\sqrt{\hbar^2+4\beta}}{2}=\frac{k}{2}\hbar \implies \hbar^2-4\beta=\hbar^2(k+1)^2$$
$$\implies \beta=\frac{\hbar^2}{4}(k^2+2k+1-1)=\hbar^2\frac{k}{2}\left(\frac{k}{2}+1\right)$$

By defining $l=k/2$ and $\alpha=m\hbar$, we get $|\beta,\alpha\rangle=\left|\hbar^2 l(l+1),m\hbar\right\rangle\equiv|l,m\rangle$, which gives us the well known quantisation rules of angular momentum:

$$\boldsymbol{L}^2\,|l,m\rangle=\hbar^2 l(l+1)\,|l,m\rangle \qquad l\in\frac{1}{2}\mathbb{N} \tag{14}$$

$$L_z\,|l,m\rangle=\hbar m\,|l,m\rangle \qquad m=-l,-l+1,\ldots,l-1,\,l \tag{15}$$

$$L_\pm\,|l,m\rangle=\hbar\sqrt{l(l+1)-m(m\pm1)}\;|l,m\rangle \tag{16}$$

Let's take a step back now. Of course, similar derivations of the quantisation rules can be found in any decent textbook on quantum mechanics [2] [3] [4], but some conceptual difficulties that have been stressed here are often left out for the reader to figure out because they are rarely of direct importance to calculations.

We will take some time to dwell on other important considerations before getting to the most important turning point in our story.

First of all, we have diagonalised angular momentum along the z axis, but as mentioned space is isotropic and therefore these quantisation rules apply for the angular momentum along any axis. For a given value of l, measurement of angular momentum of a system along any axis will always give a value of m in $-l,-l+1,...,l$. L_z is therefore just a stand-in for "angular momentum measured about a given axis, which we call z for convenience". In this context L_x and L_y can be seen as representing how rotations along the x and y axis would mix the eigenstates of L_z. To study a second measurement about, say, the x axis, it would be necessary to diagonalize L_x and express the L_z eigenstate that we got after the first measurement as a superposition of L_x eigenstates, which would give us the probability to measure the different values of m_x.

It can in fact be shown that if $|l,m\rangle_{\boldsymbol{n}}$ is an eigenstate of $\boldsymbol{n}\cdot\boldsymbol{L}$ of eigenvalue m, and the system undergoes a rotation of the coordinate axes, the rotated ket $|l,m\rangle_{\boldsymbol{n'}}$ will be an eigenstate of $\boldsymbol{n'}\cdot\boldsymbol{L}$ (where $\boldsymbol{n'}$ is the unit vector $\boldsymbol{n}$ after the rotation) with the same eigenvalue m.

Another thing we can notice is that for a given value of l, any rotation only mixes states of different values of m, but of the same l. We say that l defines a *subspace* of the algebra which is *closed under rotation*, in mathematical terms, each value of l defines a *representation* of the algebra of rotations. This means that even though angular momentum always has the same behaviour (i.e. the same commutation relations), systems with different values of l will have very different properties. This partly explains why different atomic orbitals have such different properties and can be cleanly split into categories[1] (s orbitals corresponding to $l=0$, p orbitals to $l=1$, d orbitals to $l=2$ etc)[2].

Finally, on a more fundamental note, it is interesting to think about where the quantisation of this familiar physical quantity really comes from, other than the mathematics. Indeed it seems weird that the classical properties of position and momentum keep their continuous nature in quantum mechanics, but not angular momentum. This can be seen with an analogy with other quantum systems, where quantisation of the energy spectrum appears when a system is confined inside a potential. In this case, it is precisely the requirements that the confinement brings that creates the quantised spectrum. We have seen that angular momentum has a deep connection to rotations, and a natural property of rotations is that you must come back where you started from at some point. This can be seen roughly as a form of "periodic boundary condition" that constrains how the rotational symmetry of a system can be described.

Basically, the eigenvalue of $\boldsymbol{L}^2$ tells you the different ways a wavefunction can behave under a rotationally invariant potential, and each of these "ways" represents a $2l+1$ dimensional space, i.e. a set of $2l+1$ wavefunction, one for each eigenvector of L_z for $-l\leqslant m\leqslant l$, that is closed under superposition. If you think of a sphere and integer values of

[1] Of course, this only accounts for the angular part of the wavefunctions, which are eigenfunctions of angular momentum in position representation called spherical harmonics, but crucially the radial part does not change the angular properties of the wavefunction.

[2] To see that an eigenfunction of angular momentum with a specific value for l can be expressed as a superposition of eigenfunctions of different m but the same l only, you can play around with this brilliant website `https://vinequai.com/sphericalharmonics`.

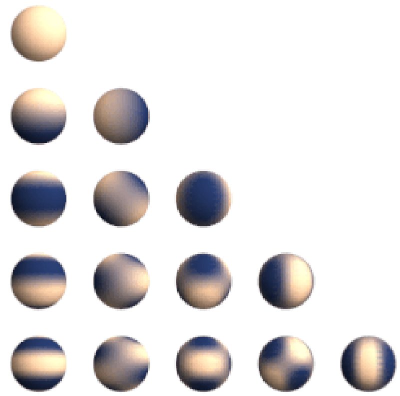

FIG. 3: Spherical harmonics for $l = 1$ to 3. The value of l (line) tells how many planes cut the sphere (blurry lines between white and blue areas) and that of m (column) how many planes contain the z axis (vertical blurry lines). Only positive values of m are displayed, as negative values are identical up to sign and complex conjugation (visually insignificant).

the eigenvalues, i.e. spherical harmonics, l tells you with how many planes you can "slice" the sphere in a symmetric way around the z axis, and m tells you how many of these planes contain the z axis.

Of course, we have only talked about integer eigenvalues. This is partly because when solving rotationally symmetric eigenfunction problems in the position representation, only those are allowed to appear in the solutions, but we have found that $l \in \frac{1}{2}\mathbb{N}$, so half integer values should also be possible[1].

In these cases, the rotational symmetry seems slightly harder to imagine, as slicing a sphere with a half integer number of planes sounds more like a riddle than a possibility. Still, we know how to relate L_z to an arbitrary rotation, no matter the value of m, so let's apply such a rotation on an eigenstate of L_z.

$$\begin{aligned} R_{\hat{z}}(\theta)\,|l,m\rangle = e^{-\frac{i}{\hbar}\theta L_z}\,|l,m\rangle &= \sum_k \frac{1}{k!}\left(\frac{i}{\hbar}\theta\right)^k L_z^k\;|l,m\rangle \\ &= \sum_k \frac{1}{k!}\left(\frac{i}{\hbar}\theta\right)^k m^k\;|l,m\rangle \\ &= e^{-\frac{i}{\hbar}\theta m}\,|l,m\rangle \end{aligned} \tag{17}$$

For systems we are accustomed to, a rotation of 2π should always bring you back to where you started, which would mean $e^{-\frac{i}{\hbar}2\pi\, m} = 1$. But for half integer values, for example $m = 1/2$, we get $e^{-\frac{i}{\hbar}2\pi\frac{1}{2}} = e^{-\frac{i}{\hbar}\pi} = -1$ [2]. So a rotation of 2π only gets you halfway to your starting point, "which is clearly nonsensical" said the physicists of the beginning of the 1920s, along with any sensible human being. Unfortunately (or rather fortunately), as those same physicists had begun to realise, the physical world seems to have decided that being sensible was far too boring.

IV. LET THERE BE SPIN

As we have already mentioned, the first attempt at a quantum theory was made by Bohr and his theory of the atom, which was later expanded on by Sommerfeld who introduced the general quantisation rules allowing for elliptical orbits:

$$\int_{Orbit} p_i\, dr_i = n_i h \qquad n_i \in \mathbb{N} \tag{18}$$

---

[1] The reason why only integer and half integer values of l are allowed comes from the topology of the Lie group of complex rotations $SU(2)$, which double covers the group of real rotations $SO(3)$, i.e. to each real rotation is associated two, and only two, complex rotations. A deeper understanding of the subject requires knowledge of group theoretic concepts which are beyond the scope of this paper.

[2] The condition for spherical harmonics to be rotationally symmetric under a rotation of 2π is precisely the reason why only integer values are allowed in this case. It is also noteworthy that a rotation of $2\pi/m$ is also equal to the identity, which accounts for the discrete rotational symmetry of the spherical harmonics seen in Figure 3.

This approach first introduced the concept of quantum number, as well as that of degeneracy of a spectrum, while predicting with good accuracy the energy levels of the atom. This is sometimes called the old quantum theory, or the semiclassical approach, and it is where our story begins.
At the time, two major problems were plaguing our understanding of the atom:

- There seemed to be exactly twice as many electronic states allowed than what was predicted, a problem referred to as *Unmechanische Zweideutigkeit* (Unmechanical doubling) or duplexity
- No more than a single electron seemed to be allowed in each of these states

Of course, there were other problems, but these ones seemed to escape any form of justification from the mechanical theories physicists had been spending centuries on.

The other great conceptual problem was the apparent quantisation of phase space brought by the Bohr-Sommerfeld quantisation rule (18), which gave the right results but appeared out of nowhere, which is why it was believed (or perhaps hoped) that a theory that would make this quantisation appear naturally would also explain these two problems.

Unfortunately, during the 1920s, such theories started to appear, championed by the likes of Schrödinger, de Broglie, Heisenberg and others, but the problems remained uncomfortably unanswered.

Even though it did not give an answer, it seemed to converge on the same problem. Indeed in 1922 the Stern Gerlach experiment was first performed. It consisted of a ray of silver atoms in an inhomogeneous magnetic field. These atoms were chosen because all their orbitals are filled, except for the last one which carries all their angular momentum (a more precise explanation of this and the experiment will be given in Section V). According to the Born-Sommerfeld theory the smallest ammount of angular momntum an electron could hold was equivalent to an $l = 1$ state[1], which would give the three possible values $m = -1, 0, 1$. In the actual experiment, only two spots appeared, which seemed to mean that m could only take two values.

This was still seen as a success for the old quantum mechanics, as it did show that some quantisation was taking place, even if it didn't exactly align with the theory. The problem is that when the full theory was derived, it turned out that $l = 0$ orbitals didn't pose a problem, and that the outer electron of the silver atom was in fact in such a state. So it carried no angular momentum and should not have been deflected at all by the apparatus, predicting a single spot!

This came back to the problem of duplexity, the electron seemed to have two possible states, and it seemed to hint that these two states had equal and opposite angular momentum.

The idea of an angular momentum of the electron had already been proposed by Ralph Kronig in 1925, who proposed that the electron could be rotating. Pauli hated the idea, stating that "it is indeed very clever but of course has nothing to do with reality", ironically enough given where our story is going. This made Kronig reluctant to publish his idea. The animosity of Pauli was not unfounded, as if the electron really was spinning, its surface would be moving much faster than the speed of light!

Note that the angular momentum of a spinning sphere of mass m is $L = 2mR^2\omega/5 = 2mRv_s/5$, where R is the radius of the sphere and v_s is the speed of a point on its surface at the equator. The measured angular momentum of the electron is $\hbar/2$, and its radius predicted by classical electrodynamics was $m = e^2/4\pi\epsilon_0 mc^2 \approx 2.8 \cdot 10^{-15} m$. This gives a speed at the equator of:

$$\frac{2}{5} mRv_s = \frac{\hbar}{2} \implies v_s = \frac{5}{4}\frac{\hbar}{mR} \approx 4.9 \cdot 10^{10} m.s^{-1} \approx 163\, c$$

More precise calculations give lower results, but a common trend is that a part of the electron is always spinning faster than light[2].

But the hints that it did have an angular momentum were stacking up, and later that same year Samuel Goudsmit and George Uhlenbeck proposed the same idea to Ehrenfest, who was much more enthusiastic than Pauli and pushed them to publish the idea. Funnily enough, they went on a trip and discussed the idea with Lorentz, who was opposed to the idea on the same basis as Pauli, so they were discouraged and asked to withdraw the publication when they returned. However Ehrenfest told them that it was too late and that it had already been sent. It is rumoured that he was so hasty with sending this publication because he had received a paper from deHaas who told him that he was

---

[1] This is because in this theory $l = 0$ would correspond to a maximally elliptic orbit, i.e. a straight line going through the nucleus, which would be impossible.
[2] Also, a spinning object can only have integer values of l, since it must return to its initial state after a rotation of 2π.

working on an experiment that could test this theory, so Ehrenfest wanted to ensure that Goudsmit and Uhlenbeck would get the credit they deserved for the idea.

The next year, in 1926, Llewellyn Thomas showed that the description of the electron carrying an angular momentum gave exactly the right results and resolved the problem of duplexity, which convinced Pauli once and for all. In 1927, he formalised the theory within the new quantum mechanics by adding two postulates:

- Particles possess an intrinsic degree of freedom, spin, which is described by an angular momentum operator called $\boldsymbol{S}$ acting on an *internal* Hilbert space $\mathcal{H}_s$, which does not correspond to the Hilbert space of the classical degrees of freedom, called the orbital space $\mathcal{H}_o$. The Hilbert space a one particle system evolves in becomes $\mathcal{H} = \mathcal{H}_o \otimes \mathcal{H}_s$.
- Two electrons can never occupy the same complete quantum state, their spin must differ if their orbital parameters are identical and the other way around.

This forms the basis of what we will study in Section V, but it is not the end of the story. Just one year later, in 1928, Paul Dirac managed to get a fully consistent relativistic equation for electrons. This was done by, loosely speaking, taking the square root of the mass shell $E^2 = p^2c^2 + m^2c^4$, and applying the canonical quantisation rules to it(we will briefly touch on this in Section VII). This made spin 1/2 naturally appear in the structure of the equation (as well as antimatter, but this is a story for another day). Better yet, the study of the Poincaré group, which is the symmetry group of space and time, makes spin appear as a fundamental parameter of these symmetries, alongside the mass.

All of this is to say that spin is not just an abstract idea or an effective model that works well, it is theoretically and experimentally unavoidable, and whether it is intuitive or not is of no concern to the universe, which seems happy to keep on going the way it is, mysterious and fascinating.

But this goes a bit too far for our purposes today, so we will focus on building a solid understanding of the theory of spin of Pauli, as a step to a better understanding of quantum mechanics as a whole.

V. PAULI'S THEORY OF SPIN

Fundamentally, Pauli's theory of spin is a non relativistic, empirical theory of spin that describes the behaviour of particles with spin, with an emphasis on spin half particles in an electromagnetic field, which represents the vast majority of the interactions we are familiar with.

This theory can be seen as arising from a simple observation: we can solve most of our problems with our description of the atom by assuming that the electron has an intrinsic angular momentum living in a 2 dimensional representation of $\mathfrak{su}(2)$, i.e. with $l = 1/2$ (according to our previous notation).

An important note is that we will almost only be speaking about spin 1/2 particles, as all the fundamental particles of matter fit in this category[1], spin 0 are of no interest to us as they transform trivially under rotation (i.e. not at all). It is in fact possible to show that higher spins can be constructed from tensor products of spin 1/2 representations[2].

We therefore define the spin operators $\boldsymbol{S}$ and $\boldsymbol{S}^2$ of respective eigenvalues $\hbar m_s$ and $\hbar^2 s(s+1)$ acting on this $2s+1$ dimesional space, the spin state space, or spin Hilbert space $\mathcal{H}_s$. This comes with all the usual properties of Hilbert spaces, in particular the closure and orthonormality relations.

$$\sum_{m_s=-s}^{s} |m_s\rangle \langle m_s| = 1_{\mathcal{H}_s} \qquad \langle m_s|m_s'\rangle = \delta_{m_s,m_s'} \tag{19}$$

Spin is therefore a dynamical quantity, whose dynamics occur in an abstract *intrinsic* space.

We now work with $s = 1/2$ and we assume that this value is fixed such that the spin of the electron, and therefore the space in which this degree of freedom evolves, becomes one of its defining properties.

Since $\boldsymbol{S}$ is an angular momentum, we can use all of our results from Section III, namely (14-16) by once again choosing z as our quantisation axis. Which allows us to write the following relations

---

[1] The bosons of the standard model are not strictly speaking "matter", hence this sentence.

[2] Formally, a spin s representation can be constructed by considering the completely symmetrised tensor product of $2s$ spin 1/2 representations. This process is completely analogous to that of addition of angular momenta (while limiting oneself to the highest spin representation). In this sense, a particle of spin s behaves exactly as a system of $2s$ spin 1/2 particles that can be found to all be aligned in a certain orientation.

$$\boldsymbol{S}^2 \left|\pm\right\rangle = \hbar^2 \frac{1}{2}\left(\frac{1}{2}+1\right)\left|\pm\right\rangle = \frac{3}{4}\hbar^2 \left|\pm\right\rangle \tag{20}$$

$$S_z \left|\pm\right\rangle = \pm\frac{\hbar}{2}\left|\pm\right\rangle \tag{21}$$

$$\begin{aligned} &S_+ \left|+\right\rangle = 0 \qquad S_+\left|-\right\rangle = \hbar\sqrt{\frac{3}{4}-\left(-\frac{1}{2}\right)\left(-\frac{1}{2}+1\right)}\left|-\right\rangle = \hbar\left|-\right\rangle \\ &S_- \left|+\right\rangle = \hbar\sqrt{\frac{3}{4}-\frac{1}{2}\left(\frac{1}{2}-1\right)}\left|+\right\rangle = \hbar\left|+\right\rangle \qquad S_-\left|-\right\rangle = 0 \end{aligned} \tag{22}$$

$$\begin{aligned} S_x\left|+\right\rangle &= \frac{1}{2}(S_+ + S_-)\left|+\right\rangle = \frac{\hbar}{2}\left|-\right\rangle \qquad & S_x\left|-\right\rangle &= \frac{1}{2}(S_+ + S_-)\left|-\right\rangle = \frac{\hbar}{2}\left|+\right\rangle \\ S_y\left|+\right\rangle &= \frac{i}{2}(S_- - S_+)\left|+\right\rangle = i\frac{\hbar}{2}\left|-\right\rangle \qquad & S_y\left|-\right\rangle &= \frac{i}{2}(S_- - S_+)\left|-\right\rangle = -i\frac{\hbar}{2}\left|+\right\rangle \end{aligned} \tag{23}$$

Where $\left|\pm\right\rangle$ are a shorthand notation for the eigenstates of S_z of respective eigenvalue $\pm\hbar/2$. It is interesting to notice that in this case for all states $m_s \neq 0$, so since the z axis is arbitrary, a measurement along any axis will always give a non zero angular momentum for the electron, which is always identical in its absolute value. This is part of what makes spin such a defining property for elementary particles.

We can easily write relations $(20-23)$ in matrix form in the basis ($\left|+\right\rangle$, $\left|-\right\rangle$):

$$\begin{aligned} &\boldsymbol{S}^2 = \frac{3}{4}\hbar^2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \qquad S_+ = \hbar\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} \qquad S_- = \hbar\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} \\ &S_x = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = \frac{\hbar}{2}\sigma_x \qquad S_y = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} = \frac{\hbar}{2}\sigma_y \qquad S_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \frac{\hbar}{2}\sigma_z \end{aligned} \tag{24}$$

The last line can be written in a compact form as:

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma} \tag{25}$$

Where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the *Pauli matrices* (get used to them, they will become our anchor to navigate this abstract space!), which have some very interesting properties. Of interest to us will be the relations[1]:

$$\sigma_i^2 = 1 \qquad \{\sigma_i, \sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 0 \quad \forall\ i \neq j \tag{26}$$

$$[\sigma_i, \sigma_j] = 2i\,\varepsilon_{ijk}\,\sigma_k \implies \sigma_i\sigma_j = i\,\varepsilon_{ijk}\,\sigma_k + \delta_{ij} \tag{27}$$

We can also deduce from (27) an identity which will be important for later. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two vectors or collections of operators which commute with all σ_i, one has:

$$\begin{aligned} (\boldsymbol{\sigma}\cdot\boldsymbol{A})(\boldsymbol{\sigma}\cdot\boldsymbol{B}) = \sigma_i A_i \sigma_j B_j &= \sigma_i\sigma_j A_i B_j \\ &= (i\,\varepsilon_{ijk}\,\sigma_k + \delta_{ij})A_i B_j \\ &= A_i B_i + i\sigma_k\,\varepsilon_{ijk}\,A_i B_j \\ &= \boldsymbol{A}\cdot\boldsymbol{B} + i\boldsymbol{\sigma}\cdot\boldsymbol{A}\times\boldsymbol{B} \end{aligned} \tag{28}$$

[1] (26) are the defining properties of the *Clifford algebras*, the study of which is intimately related to spin.

With the properties $(26-27)$, one can actually show that any polynomial in $\boldsymbol{\sigma}$ can be written as a linear term in $\boldsymbol{\sigma}$ plus a constant, (28) only being one specific useful example of this property.

Coming back to (25), this allows us to write the spin operator along an axis given by a unit vector $\boldsymbol{n} = (\sin\theta\cos\varphi\,,\ \sin\theta\sin\varphi\,,\ \cos\theta)$, where θ and φ are the polar angles of spherical coordinates, as:

$$\boldsymbol{n}\cdot\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma} = \frac{\hbar}{2}\begin{pmatrix}\cos\theta & \sin\theta e^{-i\varphi}\\ \sin\theta e^{i\varphi} & -\cos\theta\end{pmatrix} \tag{29}$$

Before going back to the orbital state space, it can be useful to notice that there exists a convenient visualisation of single particle spin 1/2 states. In all generality, we can write $|\psi\rangle = a\,|+\rangle + b\,|-\rangle$ where $a, b \in \mathbb{C}$. Enforcing normalisation reduces the number of independent real parameters from 4 to 3. Writing $a = c + id$ and $b = e + if$ with $c, d, e, f \in \mathbb{R}$, we get the constraint:

$$|a|^2 + |b|^2 = |c|^2 + |d|^2 + |e|^2 + |f|^2 = 1 \tag{30}$$

This gives us the equation for the *3-sphere*, i.e. a unit sphere living in a 4 dimensional space whose surface is 3 dimensional. We can write the state in a polar form by defining $a = |a|e^{i\alpha}$ and $b = |b|e^{i\beta}$. Using (30), we can write without loss of generality:

$$|\psi\rangle = e^{i\alpha}\cos\gamma\,|+\rangle + e^{i\beta}\sin\gamma\,|-\rangle \tag{31}$$

Furthermore, the condition $|a|, |b| \geq 0$ implies $\cos\gamma, \sin\gamma \geq 0 \implies \gamma \in [0, \pi/2]$. We can therefore define $\gamma = \theta/2$, $\theta \in [0, \pi]$. Also, since two state with a differing global phase are physically identical, we can factor out and omit $e^{i\alpha}$ and define $\varphi = \beta - \alpha$, $\phi \in [0, 2\pi]$, which finally givess us:

$$|\psi\rangle = \cos\frac{\theta}{2}\,|+\rangle + e^{i\varphi}\sin\frac{\theta}{2}\,|-\rangle \tag{32}$$

We have therefore parametrised our state completely with respect to two polar angles ϕ and θ, which can be represented on a familiar sphere in 3 dimensions. This is the *Bloch sphere* (see Figure 4). This representation is extremely useful as it not only allows us to easily visualise a spin state, but also has very interesting properties with respect to the expectation value of the Pauli matrices, and therefore the spin operators for spin 1/2 particles. By computing these expectation values with (32), we get:

$$\begin{aligned}\langle\boldsymbol{\sigma}\rangle &= \left(\cos\frac{\theta}{2}\,\langle+| + e^{-i\varphi}\sin\frac{\theta}{2}\,\langle-|\right)\boldsymbol{\sigma}\left(\cos\frac{\theta}{2}\,|+\rangle + e^{i\varphi}\sin\frac{\theta}{2}\,|-\rangle\right)\\ &= \cos^2\frac{\theta}{2}\,\langle+|\,\boldsymbol{\sigma}\,|+\rangle + \sin^2\frac{\theta}{2}\,\langle-|\,\boldsymbol{\sigma}\,|-\rangle + \frac{1}{2}\sin(\theta)(e^{-i\varphi}\,\langle-|\,\boldsymbol{\sigma}\,|+\rangle + e^{i\varphi}\,\langle+|\,\boldsymbol{\sigma}\,|-\rangle)\\ &= \begin{pmatrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\\ \cos\theta\end{pmatrix} = \boldsymbol{n}\end{aligned}$$

The expectation value of the spin along any axis is therefore entirely defined by the projection of a unit vector $\boldsymbol{n}$ parametrised by the angles θ and φ of (32) along this axis (up to a factor of $\hbar/2$). This makes the Bloch sphere incredibly useful when dealing with complicated calculations depending on spin orientations, which is exactly why the main use of it is in quantum computing applications[1].

Now that we have a basic idea of how the spin state space works, or at least of the tools that are available to us, we can think about how it fits with the original formulation of quantum mechanics. The "spinless" state space is often called the *orbital* space $\mathcal{H}_o$, as it deals with the parameters set by the orbital motion of the particles. Our new, total state space therefore becomes the tensor product of the spin and orbital spaces: $\mathcal{H} = \mathcal{H}_o \otimes \mathcal{H}_s$.

---

[1] Points inside the sphere, for which $|a|^2 + |b|^2 < 1$, are also used. They represent *mixed states*, which are systems in a statistical mixture or states. This means that an additional, non quantum uncertainty applies to the system, as in statistical mechanics. It can also stand for an entangled state, for which the additional uncertainty is purely quantum and linked to the treatment of systems of several particles in quantum mechanics, for which all the information on each constituent separately is not equivalent to the entire information on the complete state. We will not go into detail about this as it is not restricted to spin and deserves a separate treatment.

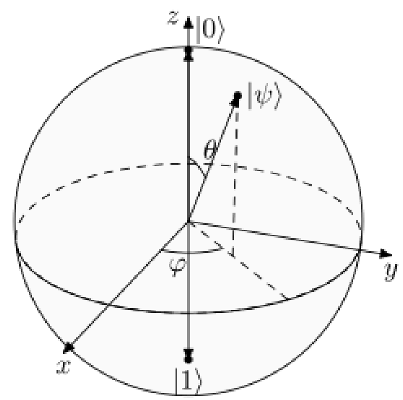

FIG. 4: The Bloch sphere representation of a single spin 1/2 particle state, with our notation $|0\rangle \equiv |+\rangle$ and $|1\rangle \equiv |-\rangle$. Courtesy of `https://en.wikipedia.org/wiki/Bloch_sphere`

A basis of $\mathcal{H}$ can therefore be written as the tensor product of a basis of $\mathcal{H}_o$ and one of $\mathcal{H}_s$. For simplicity, we will work with the position basis, $|\boldsymbol{r}\rangle$, in $\mathcal{H}_o$ and the basis defined by the eigenkets of S_z, $|\pm\rangle$, in $\mathcal{H}_s$. We therefore write a basis element of $\mathcal{H}$ as $|\boldsymbol{r},\pm\rangle \equiv |\boldsymbol{r}\rangle \otimes |\pm\rangle$.

This allows for the creation of a *spin-position representation* as it is possible to define a closure relation by:

$$\sum_{\pm}\int_{\mathbb{R}^3} d\boldsymbol{r}\ |\boldsymbol{r},\pm\rangle\langle\boldsymbol{r},\pm| = \sum_{\pm}|\pm\rangle\langle\pm|\otimes\int_{\mathbb{R}^3} d\boldsymbol{r}\ |\boldsymbol{r}\rangle\langle\boldsymbol{r}| = 1_{\mathcal{H}_s}\otimes 1_{\mathcal{H}_o} = 1_{\mathcal{H}} \equiv 1 \tag{33}$$

We can also define the spin dependent wavefunction $\langle\boldsymbol{r},\pm|\psi\rangle = \psi_\pm(\boldsymbol{r})$ and decompose $|\psi\rangle$ and $\langle\psi|$ in this new basis.

$$\begin{aligned} |\psi\rangle &= \sum_{\pm}\int_{\mathbb{R}^3} d\boldsymbol{r}\ |\boldsymbol{r},\pm\rangle\langle\boldsymbol{r},\pm|\psi\rangle = \int_{\mathbb{R}^3} d\boldsymbol{r}\,(\psi_+(\boldsymbol{r})\,|\boldsymbol{r},+\rangle + \psi_-(\boldsymbol{r})\,|\boldsymbol{r},-\rangle\,) \\ &\implies \langle\psi| = \int_{\mathbb{R}^3} d\boldsymbol{r}\,(\psi_+^*(\boldsymbol{r})\,\langle\boldsymbol{r},+| + \psi_-^*(\boldsymbol{r})\,\langle\boldsymbol{r},-|\,) \end{aligned} \tag{34}$$

Thanks to this decomposition, and by remembering that the spin operators only act on $\mathcal{H}_s$, we can study their action on $|\psi\rangle$.

$$\begin{aligned} S_z|\psi\rangle &= \sum_{\pm}\int_{\mathbb{R}^3} d\boldsymbol{r}\,\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}|\boldsymbol{r},\pm\rangle\langle\boldsymbol{r},\pm|\psi\rangle = \int_{\mathbb{R}^3} d\boldsymbol{r}\,\left(\frac{\hbar}{2}\psi_+(\boldsymbol{r})\,|\boldsymbol{r},+\rangle - \frac{\hbar}{2}\psi_-(\boldsymbol{r})\,|\boldsymbol{r},-\rangle\right) \\ S_x|\psi\rangle &= \int_{\mathbb{R}^3} d\boldsymbol{r}\,\left(\frac{\hbar}{2}\psi_-(\boldsymbol{r})\,|\boldsymbol{r},+\rangle + \frac{\hbar}{2}\psi_+(\boldsymbol{r})\,|\boldsymbol{r},-\rangle\right) \\ S_y|\psi\rangle &= \int_{\mathbb{R}^3} d\boldsymbol{r}\,\left(i\frac{\hbar}{2}\psi_-(\boldsymbol{r})\,|\boldsymbol{r},+\rangle - i\frac{\hbar}{2}\psi_+(\boldsymbol{r})\,|\boldsymbol{r},-\rangle\right) \end{aligned} \tag{35}$$

We can see that the wavefunctions $\psi_\pm(\boldsymbol{r})$ act like components of a two component object that transforms under the application of the spin operators like a component vector acted on by a matrix. We write this object and its conjugate as:

$$\boldsymbol{\psi}(\boldsymbol{r}) = \begin{bmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{bmatrix} \qquad \boldsymbol{\psi}^\dagger(\boldsymbol{r}) = \begin{bmatrix}\psi_+^*(\boldsymbol{r}) & \psi_-^*(\boldsymbol{r})\end{bmatrix} \tag{36}$$

We can therefore rewrite (35) as:

$$S_z\boldsymbol{\psi}(\boldsymbol{r}) = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{bmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix}\psi_+(\boldsymbol{r})\\ -\psi_-(\boldsymbol{r})\end{bmatrix}$$

$$S_x\boldsymbol{\psi}(\boldsymbol{r}) = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{bmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix}\psi_-(\boldsymbol{r})\\ \psi_+(\boldsymbol{r})\end{bmatrix} \tag{37}$$

$$S_y\boldsymbol{\psi}(\boldsymbol{r}) = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{bmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{bmatrix} = i\frac{\hbar}{2}\begin{bmatrix}-\psi_-(\boldsymbol{r})\\ \psi_+(\boldsymbol{r})\end{bmatrix}$$

(36) also allows us to write scalar products in the following compact form:

$$\langle\phi|\psi\rangle = \sum_{\pm}\int_{\mathbb{R}^3} d\boldsymbol{r}\ \langle\phi|\boldsymbol{r},\pm\rangle\,\langle\boldsymbol{r},\pm|\psi\rangle = \int_{\mathbb{R}^3} d\boldsymbol{r}\,(\phi_+^*(\boldsymbol{r})\psi_+(\boldsymbol{r}) + \phi_-^*(\boldsymbol{r})\psi_-(\boldsymbol{r})\,) = \int_{\mathbb{R}^3} d\boldsymbol{r}\,\boldsymbol{\phi}^\dagger(\boldsymbol{r})\,\boldsymbol{\psi}(\boldsymbol{r}) \tag{38}$$

We can now develop an interpretation for the quantities $\psi_\pm(\boldsymbol{r})$ by enforcing normalisation of $|\psi\rangle$, as we are accustomed to.

$$\langle\psi|\psi\rangle = \int_{\mathbb{R}^3} d\boldsymbol{r}\,\boldsymbol{\psi}^\dagger(\boldsymbol{r})\,\boldsymbol{\psi}(\boldsymbol{r}) = \int_{\mathbb{R}^3} d\boldsymbol{r}\,(\,|\psi_+(\boldsymbol{r})|^2 + |\psi_-(\boldsymbol{r}|)^2\,) = 1 \tag{39}$$

$\left|\psi_{+(-)}(\boldsymbol{r})\right|^2$ is therefore the probability density of finding the electron in its spin state $+$ $(-)$ in the region $\boldsymbol{r}+d\boldsymbol{r}$.

So $\boldsymbol{\psi}(\boldsymbol{r})$ is a useful tool for calculations such as (38-39), but it is not just that. To see this, we will study the behaviour of an electron state under a rotation and notice that $\boldsymbol{\psi}(\boldsymbol{r})$ truly is the most natural object to use in the case of a spin 1/2 particle.

If we rotate the coordinate axes, there will be a rotation in both orbital and spin state spaces. Since these spaces are independant, the total rotation can be decomposed as the product of a rotation on $\mathcal{H}_o$ and a rotation on $\mathcal{H}_s$, i.e. $R_{\boldsymbol{n}}(\theta) = R^o_{\boldsymbol{n}}(\theta)R^s_{\boldsymbol{n}}(\theta)$. Since both spaces have an associated angular momentum L and S, we can use (4) to express these rotation operators. Moreover, since $\boldsymbol{L}$ and $\boldsymbol{S}$ commute, we have[1]:

$$R_{\boldsymbol{n}}(\theta) = e^{-\frac{i}{\hbar}\theta\boldsymbol{n}\cdot\boldsymbol{L}}e^{-\frac{i}{\hbar}\theta\boldsymbol{n}\cdot\boldsymbol{S}} = e^{-\frac{i}{\hbar}\theta\boldsymbol{n}\cdot(\boldsymbol{L}+\boldsymbol{S})} \equiv e^{-\frac{i}{\hbar}\theta\boldsymbol{n}\cdot\boldsymbol{J}} \tag{40}$$

Where $\boldsymbol{J} \equiv \boldsymbol{L}+\boldsymbol{S}$ is the total angular momentum of the particle. We can show that $\boldsymbol{J}$ is indeed an angular momentum by checking its commutation relations.

$$\begin{aligned}
[J_i, J_j] &= (L_i+S_i)(L_j+S_j) - (L_j+S_j)(L_i+S_i)\\
&= L_iL_j + S_iS_j + L_iS_j + S_iL_j - L_jL_i - S_jS_i - L_jS_i - S_jL_i\\
&= [L_i, L_j] + [S_i, S_j] + L_iS_j - L_iS_j + L_jS_i - L_jS_i\\
&= i\hbar\,\varepsilon_{ijk}\,L_k + i\hbar\,\epsilon_{ijk}\,S_k\\
&= i\hbar\,\varepsilon_{ijk}\,(L_k + S_k)\\
&= i\hbar\,\varepsilon_{ijk}\,J_k
\end{aligned}$$

This quantity will not be of great interest to us right now but is one of the most useful objects when it comes to the study of the energy levels of the atom and the general dynamics of spin 1/2 particles.

We focus on the rotation of the spin coordinates $R^s_{\boldsymbol{n}}(\theta)$. To this end, we can notice that (28) implies that for $\boldsymbol{n}$ a unit vector:

$$\begin{aligned}
(\boldsymbol{n}\cdot\boldsymbol{\sigma})^2 &= \boldsymbol{n}\cdot\boldsymbol{n} + i\boldsymbol{\sigma}\cdot\boldsymbol{n}\times\boldsymbol{n} = \boldsymbol{n}^2 = 1\\
\implies (\boldsymbol{n}\cdot\boldsymbol{\sigma})^{2k} &= 1 \qquad (\boldsymbol{n}\cdot\boldsymbol{\sigma})^{2k+1} = \boldsymbol{n}\cdot\boldsymbol{\sigma}
\end{aligned} \tag{41}$$

---

[1] The general formula for the product of exponentials of operators is given as an exponential of an infinite series of operators by the Baker-Campbell-Hausdorff formula as $e^Ae^B = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}(\,[A,[A,B]]-[B,[A,B]]\,)+\cdots}$.

This and (25) allows us to rewrite the rotation operator as:

$$\begin{aligned} R^s_{\boldsymbol{n}}(\theta) = e^{-\frac{i}{\hbar}\theta\,\boldsymbol{n}\cdot\boldsymbol{S}} &= e^{-i\frac{\theta}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma}} \\ &= \sum_k \frac{1}{k!}\left(-i\frac{\theta}{2}\right)^k (\boldsymbol{n}\cdot\boldsymbol{\sigma})^k \\ &= \sum_k \frac{(-1)^k}{(2k)!}\left(\frac{\theta}{2}\right)^{2k} - i\,\boldsymbol{n}\cdot\boldsymbol{\sigma}\sum_k \frac{(-1)^k}{(2k+1)!}\left(\frac{\theta}{2}\right)^{2k+1} \\ &= \cos\left(\frac{\theta}{2}\right) - i\,\boldsymbol{n}\cdot\boldsymbol{\sigma}\sin\left(\frac{\theta}{2}\right) \end{aligned} \tag{42}$$

We find in a more general way the statement at the end of Section III that $R^s_{\boldsymbol{n}}(2\pi) = -1$. Since the object $\boldsymbol{\psi}(\boldsymbol{r})$ transforms according to this rotation law, it is clear that it cannot be a vector, in the sense that it cannot be represented as a simple arrow in $3d$ space. As it describes the behaviour of spin 1/2 particles, $\boldsymbol{\psi}(\boldsymbol{r})$ is called a *spinor*[1].

These objects are sometimes called the "square roots of vectors" as they rotate half as fast as classical vectors. This analogy is actually much more far reaching than one might think, as we will briefly see in Section VII. It has its roots in pure mathematics, but as the stubborn physicists we are, "spins half as fast as a vector" will be justification enough.

To get the formula for the transformation of a spinor after a rotation, we can use the fact that any operator of the orbital space acts on $\psi_+(\boldsymbol{r})$ and $\psi_-(\boldsymbol{r})$ equally, i.e. it is proportional to the identity matrix. Using (2) and (42), we get:

$$R_{\boldsymbol{n}}(\theta)\boldsymbol{\psi}(\boldsymbol{r}) = R^s_{\boldsymbol{n}}(\theta)R^o_{\boldsymbol{n}}(\theta)\boldsymbol{\psi}(\boldsymbol{r}) = \left(\cos\left(\frac{\theta}{2}\right) - i\,\boldsymbol{n}\cdot\boldsymbol{\sigma}\sin\left(\frac{\theta}{2}\right)\right)\boldsymbol{\psi}(\mathcal{R}_{\boldsymbol{n}}(-\theta)\boldsymbol{r}) \tag{43}$$

Where $\mathcal{R}_{\boldsymbol{n}}(\theta)$ is the usual $3d$ rotation matrix which can be found using Rodrigue's formula.

The fact that the wavefunction takes on a minus sign after a full rotation does seem peculiar, but as it only corresponds to a global phase factor, it has no impact on physical measurements of an observable for a single spinor wavefunction. So how do we know it exists? Because it actually has an impact when it comes to interference in a system of spinor wavefunctions, which is a crucial part of experimental methods like neutron interferometry, but we will not have time to study these kinds of phenomena here.

Now that we have studied the kinematics of spinors, let's get into their dynamics. The natural question to ask would be how they interact with an electromagnetic field. The coupling with the electric field presents no difficulty, as we are studying simple charged particles, and their behaviour under rotation has no impact on this interaction, the potential energy associated to the electric field is therefore $V_e = q\phi$, where q is the charge of the particle and ϕ the electric potential.

What the behaviour under rotation does impact though, is the coupling with the magnetic field. It can be shown classically that an object of charge q and mass m rotating with an angular momentum $\boldsymbol{L}$ creates a dipolar magnetic field (basically a tiny magnet) which interacts with a magnetic field $\boldsymbol{B}$ with an interaction energy $V_m = -\boldsymbol{\mu}\cdot\boldsymbol{B}$ [5] where $\boldsymbol{\mu} = q\boldsymbol{L}/2m$ is the dipole moment of the object. This means that a magnetic dipole tends to both align with the surrounding magnetic field and to go to regions of higher field intensity, or lower intensity while it is anti-aligned (which is indeed why the Stern-Gerlach apparatus needed an inhomogeneous magnetic field to work).

However, due to relativistic effects we will talk about in Section VII, the actual dipole moment of the electron is twice as large as the classical prediction, which gives $\boldsymbol{\mu}_e = 2q\boldsymbol{S}/2m = q\hbar\boldsymbol{\sigma}/2m$[2].

Finally, the interaction between particles and fields require the substitution of the momentum operator $\boldsymbol{P} \to \boldsymbol{\Pi} = \boldsymbol{P} - q\boldsymbol{A}$[3], where $\boldsymbol{A}$ is the vector potential defined by $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$. This can be derived classically through the Lagrangian of a particle in an electromagnetic field (this should be included in any standard textbook on analytical mechanics), or by enforcing gauge invariance of the Hamiltonian in a relativistic theory.

---

[1] The fact that the wavefunction of a spinor takes on a minus sign after a complete rotation is linked to the topology of $SU(2)$, i.e. the group of rotations in which it is possible to have half integer eigenvalues of the angular momentum. It can be shown that the behaviour of a spinor under a rotation actually depends on the path a unit vector defining the rotation of the axes would take during this transformation, and not just its beginning and end points (unlike a vector). The interested reader may want to learn about Berry phases.

[2] The actual value of the ratio between the actual dipole moment and the classical prediction is called the g factor and is slightly bigger than 2 due to higher order effects, which we will not consider here. Experimental data shows $g = 2.00231930436092(36)$.

[3] The kinetic part of the dipole field interaction $q\boldsymbol{L}/2m$ is actually included in this term, which becomes clear once it is Taylor expanded as $\frac{1}{2m}(\boldsymbol{P} - q\boldsymbol{A})^2 \approx \frac{\boldsymbol{P}^2}{2m} - \frac{q}{2m}\boldsymbol{L}\cdot\boldsymbol{B} + \ldots$. The Spin interaction term must be added precisely because it does not arise from kinetic considerations.

We can therefore write the full non relativistic Hamiltonian for a spin 1/2 particle in an electromagnetic field, which is called the Pauli Hamiltonian:

$$H = \frac{(\boldsymbol{P} - q\boldsymbol{A})^2}{2m} - \frac{q\hbar}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B} + q\phi \tag{44}$$

A final thing to notice is:

$$\begin{aligned}(\boldsymbol{\Pi}\times\boldsymbol{\Pi})\psi &= (\boldsymbol{P} - q\boldsymbol{A})\times(\boldsymbol{P} - q\boldsymbol{A})\psi \\ &= -q(\boldsymbol{P}\times(\boldsymbol{A}\psi) + \boldsymbol{A}\times(\boldsymbol{P}\psi)) \\ &= -q((\boldsymbol{P}\times\boldsymbol{A})\psi + (\boldsymbol{P}\psi)\times\boldsymbol{A} + \boldsymbol{A}\times(\boldsymbol{P}\psi)) \\ &= -q(-i\hbar(\boldsymbol{\nabla}\times\boldsymbol{A})\psi + (\boldsymbol{P}\psi)\times\boldsymbol{A} - (\boldsymbol{P}\psi)\times\boldsymbol{A}) \\ &= iq\hbar\boldsymbol{B}\psi\end{aligned}$$

$$\implies \boldsymbol{\Pi}\times\boldsymbol{\Pi} = iq\hbar\boldsymbol{B} \tag{45}$$

Which, using (28), allows us to write the Pauli Hamiltonian in its standard form:

$$H = \frac{1}{2m}\left(\boldsymbol{\Pi}\cdot\boldsymbol{\Pi} + i\boldsymbol{\sigma}\cdot(iq\hbar\boldsymbol{B})\right) + q\phi = \frac{1}{2m}\left(\boldsymbol{\Pi}\cdot\boldsymbol{\Pi} + i\boldsymbol{\sigma}\cdot\boldsymbol{\Pi}\times\boldsymbol{\Pi}\right) + q\phi$$

$$\implies H = \frac{(\boldsymbol{\sigma}\cdot(\boldsymbol{P} - q\boldsymbol{A}))^2}{2m} + q\phi \tag{46}$$

We now have a fully dynamical model of the non relativistic electron in an electromagnetic field!

It is possible to find that Ehrenfest's theorem applied to the non spin part of (44) gives back the Lorentz "force" [6][1].

In addition to this the spin part gives very interesting behaviour. If we first assume that there is only a homogeneous magnetic field $\boldsymbol{B} = B_0\boldsymbol{n}$ so that the wavefunction is space independent, we can essentially forget about the spatial part of the Hamiltonian and only focus on the spin part. In this case writing the Schrödinger equation for a pure spin state gives:

$$i\hbar\partial_t\boldsymbol{\psi} = -\frac{q\hbar}{2m}B_0\boldsymbol{n}\cdot\boldsymbol{\sigma}\boldsymbol{\psi} \implies \delta\boldsymbol{\psi} \approx i\delta t\frac{q}{\hbar m}B_0\,\boldsymbol{n}\cdot\boldsymbol{S}\boldsymbol{\psi}$$

$$\implies \boldsymbol{\psi}(t) = e^{i\frac{qB_0}{m}t\frac{\boldsymbol{n}\cdot\boldsymbol{S}}{\hbar}}\boldsymbol{\psi}(0) = R^s_{\boldsymbol{n}}\left(\frac{qB_0}{m}t\right)\boldsymbol{\psi}(0)$$

Where a reasoning similar to (2) and (4) has been used. This means that our original spinor *precesses* in spin space with an angular frequency qB_0/m, the *Larmor frequency*, around the axis given by $\boldsymbol{n}$. Once again, the Bloch sphere is very useful for this as this precession is exactly equivalent to the precession of the state vector around the unit vector $\boldsymbol{n}$ on the Bloch sphere.

Finally, it is possible to study the Stern-Gerlach experiment with this formalism. The precise description is rather involved, so we will only give a handwavy explanation of the phenomenon.

In the experiment, silver atoms are deflected in an inhomogeneous magnetic field. Since the atoms are neutral and no electric field is present, both the electrical potential and the minimal coupling vanish. Since all the atomic orbitals are filled except for the last one, all the spins of the inner electrons cancel out, and the entire spin of the atom is carried by the outermost electron[2].

This is because to measure the total angular momentum we must use the operator $\boldsymbol{J} = \sum_i \boldsymbol{L}_i + \sum_i \boldsymbol{S}_i$, where $\boldsymbol{L}_i$ and $\boldsymbol{S}_i$ are the orbital angular momentum and spin operators for particle i. Since the shells are full, the eigenvalue of the orbital angular momentum part is a sum of terms proportional to $\sum_{m=-l}^{l} m = 0$, except for the last orbital,

---

[1] The derivation is not very complicated but slightly lengthy, and is therefore left to the reader, which will find helpful to recall that for $A(\boldsymbol{R},\boldsymbol{t})$ an operator that does not depend on $\boldsymbol{P}$, one has $[P_i, A] = -i\hbar\partial_i A$.

[2] In (44), it is not straightforward to relate the charge q to the charges of our problem, as it is the equation for a single electron. We assume that all the charge terms can be taken to be 0, except in the spin-magnetic field interaction, where it represents the charge of the electron $-e$, as it is it that interacts with the field.

which turns out to be an $l = 0$ orbital, so it doesn't contribute. For the spin part, each orbital except for the last is made up of two electrons, one of spin eigenvalue $\hbar/2$ and the other of eigenvalue $-\hbar/2$(because of the Fermi exclusion principle which we will discuss in Section VI), so the sum of the spins gives $\hbar/2 - \hbar/2 = 0$.

We assume that the magnetic field presents a gradient in the z direction. Using Ehrenfest's theorem, one gets:

$$\frac{d}{dt}\langle P_z\rangle = \frac{1}{i\hbar}\langle [P_z, H]\rangle = \frac{1}{i\hbar}\left\langle \left[P_z, \frac{\boldsymbol{P}^2}{2m} + \frac{e\hbar}{2m}B_z(z)\sigma_z\right]\right\rangle = \frac{1}{i\hbar}\frac{e\hbar}{2m}\langle \sigma_z [P_z, B_z(z)]\rangle = -\frac{e\hbar}{2m}\partial_z B_z \langle \sigma_z\rangle$$

Since for a state $|\pm\rangle$, $\langle\sigma_z\rangle = \pm 1$, atoms with an S_z eigenvalue of $\hbar/2$ will be deflected toward the areas of lower intensity of the magnetic field, and on the contrary atoms with an eigenvalue $-\hbar/2$ will be deflected toward areas of higher intensity. A beam of atoms will therefore create two distinct spots[1]! We can also see that if we had used an electron the mass would have been much smaller, making the deflection much more extreme and not easily detectable, and since it carries a charge it would have felt the Lorentz force, which is much more intense and would have completely hidden the phenomenon.

VI. THE SYMMETRISATION POSTULATE AND SPIN STATISTICS

We have spent the last section discussing the kinematics and dynamics of particles with spin, in particular their description as spinors. In this section, we will deal with some more far reaching consequences of the properties of spin. To do this, oddly enough, we will not be talking that much about spin for a while.

A fundamental fact of quantum mechanics is that identical particles are indistinguishable. This means that for a system of two electrons, for example, it is meaningless to assign one wavefunction to a particular electron and one to the other. There must still be two wavefunctions in our description of the system though, as a measurement will involve two particles, which cannot arise from a single wavefunction. The complete Hilbert space of a system of N particles is therefore the tensor product of the N Hilbert spaces of the particles.

However, all states in this new Hilbert space do not represent physical states. This comes back to our criterion of indistinguishability of particles. Consider a system of two particles, we denote ξ_i the set of parameters that characterises particles i (including its spin eigenvalue), and $|\cdot\rangle_i$ its state in its original single particle Hilbert space (that, once again, is identical for both particles). We can write $|\xi_1\rangle_1 \otimes |\xi_2\rangle_2 \equiv |\xi_1, \xi_2\rangle$. According to the principle of indistinguishability of particles, the states $|\xi_1, \xi_2\rangle$ and $|\xi_2, \xi_1\rangle$ are physically equivalent[2].

As we have seen though, physically equivalent states can differ by a seemingly insignificant global phase, which means we can write $|\xi_1, \xi_2\rangle = e^{i\alpha}|\xi_2, \xi_1\rangle$[3]. But repeating this swap of particles gives us our original state back, therefore $e^{2i\alpha} = 1 \implies e^{i\alpha} = \pm 1$. We say that if $|\xi_1, \xi_2\rangle = |\xi_2, \xi_1\rangle$ the state is symmetric, and if $|\xi_1, \xi_2\rangle = -|\xi_2, \xi_1\rangle$ it is antisymmetric. This seemingly insignificant fact is one of the core tenets keeping our universe together, and we will understand why shortly.

The first consequence of this is that all the states that a single particle can occupy in its original Hilbert space must have the same symmetry, else if it was in a superposition of these states, swapping the particles would give a state that is neither symmetrical nor antisymmetrical. Let ξ^a the parameters of a symmetric state and ξ^b that of an antisymmetric, if they were both physically possible states, we would have $|\xi_1, \xi_2\rangle = |\xi_1^a, \xi_2\rangle + |\xi_1^b, \xi_2\rangle = |\xi_2, \xi_1^a\rangle - |\xi_2, \xi_1^b\rangle \neq |\xi_2, \xi_1\rangle$.

This can also be generalised to a system of N identical particles. The process of exchanging two identical particles is called a *transposition*, while the general exchange of any number of particles is called a *permutation*. A permutation can always be written as the non commutative product of transpositions. If a permutation P requires an even number of transpositions, it is called even, and if it requires an odd number of transpositions, it is called odd.

If a state $|\psi\rangle$ is such that $P|\psi\rangle = |\psi\rangle$ for any permutation, it is called *completely symmetric*. On the contrary, if $|\psi\rangle$ is such that $P|\psi\rangle = \varepsilon|\psi\rangle$ with $\varepsilon = -1$ for even permutations and $\varepsilon = +1$ for even permutation, it is said to be *completely antisymmetric*.

Let's show that this property of symmetry and antisymmetry of identical particles must hold for any number of particles by considering a system of 3 identical particles, and finding it intuitive that it should work for any number.

---

[1] A more precise treatment of the problem shows that for an arbitrary spinor, the two components evolve independently and do indeed create two spots [7].

[2] Notice that it is completely legitimate to exchange the parameters of the two particles, as their parameter space, or the set of parameters they are allowed to take, is identical.

[3] We are acutally making an assumption here, which boils down to admitting the symmetrisation postulate that we will state shortly. It is in fact possible that states of swapped identical particles differ by a unitary operator as long as it commutes with all observables, this is why the symmetrisation postulate is a *postulate*, it rules out this possibility by assumption. The formalism in terms of these operators has been left out of this paper, as it is rather cumbersome and has for sole purpose to be ruled out by a postulate and proven false by relativistic theories.

Let $|\xi_1, \xi_2, \xi_3\rangle$ be a state that is symmetrical in the exchange of all particles, except for that of the exchange of the particles in the second and third position of the ket. Therefore, we have:

$$|\xi_1, \xi_2, \xi_3\rangle = -|\xi_1, \xi_3, \xi_2\rangle \quad \text{and} \quad |\xi_1, \xi_2, \xi_3\rangle = |\xi_2, \xi_1, \xi_3\rangle = |\xi_3, \xi_1, \xi_2\rangle = |\xi_1, \xi_3, \xi_2\rangle$$
$$\implies |\xi_1, \xi_2, \xi_3\rangle = -|\xi_1, \xi_3, \xi_2\rangle = |\xi_1, \xi_3, \xi_2\rangle = 0$$

Therefore this state is not physically possible!

It can be seen that the symmetry property is in fact not a property of the state, but of the particles themselves!

We arrive at the following general statement, which is referred to as the *symmetrisation postulate*:

- The state space of a system of identical particles is limited to the set of states that are either completely symmetric or completely antisymmetric with respect to all permutations. The particles whose state space is completely symmetric are called *Bosons* and the particles whose state space is completely antisymmetric are called *Fermions*.

This is, for one, incredible. We have found that the universe is made up of two fundamentally different types of particles by thinking about what it means for two of them to be swapped!

And it's not all, because now is when we get back to spin. A fact that is purely empirical in non relativistic quantum mechanics is that all integer spin particles are Bosons and all half integer spin particles are Fermions. This holds both for fundamental and composite particles[1]. It can be proven in relativistic quantum theories, as we will discuss in Section VII, but it took Pauli until 1940 to discover this fact, even though his exclusion principle was first proposed in 1925.

But we still haven't found how to construct a physical state for these two types of particles, nor have we really thought about their properties.

Let's start with Bosons. If two particles which are in the same one-particle state, with the set of parameters ξ^a, the state is unchanged under permutation, $|\xi_1^a, \xi_2^a\rangle = |\xi_2^a, \xi_1^a\rangle$. Since $\xi_1^a = \xi_2^a$ the numbering of the parameter set of identical one-particle states is superfluous, we will now only write $|n_a\rangle$ or $|n_a, n_b\rangle$ or, in general, $|n_1, n_2, ...\rangle$, where n_i denotes the number of particles in the one-particle state i, this is called the *occupation number formalism*.

We consider M different one-particle states occupied by $N = \sum_{i=1}^N n_i$ particles. There are $N!$ possible permutations of the particles. But there are n_i different particles in each one-particle state i, which makes $n_i!$ possible permutations of these particles, which are overcounted in $N!$ since they are exactly the same many particle states before and after permutation. There are therefore $N!/n_1!n_2!...n_M!$ different distinct possible permutations.

A many particle Boson state must be completely symmetric with respect to permutations, the easiest way to achieve this is to simply add all distinct permutations of the many particle states of our system and to normalise this "symmetrised state". This can be written:

$$|\psi\rangle_{Boson} = \left(\frac{n_1!n_2!...n_M!}{N!}\right)^{1/2} \sum_\alpha P_\alpha |n_1, n_2, ..., n_M\rangle \tag{47}$$

Where alpha is the set of all distinct permutations.

For Fermions, the situation is slightly different. To see why, as we did for Bosons we consider a state of two Fermions in the same one-particle state $|\xi_1^a, \xi_2^a\rangle = -|\xi_2^a, \xi_1^a\rangle$. But since $\xi_1^a = \xi_2^a$, since both are parameters of the exact same state, this state is equal to its opposite and must therefore vanish.

Stated differently, two Fermions cannot occupy the same state. We just derived Fermi's exclusion principle from the symmetrisation postulate!

The reason why there can be only two electrons in each atomic orbital is because there can be only one electron per quantum state, and since the spin of the electron is 1/2, it can only be in two different states for each orbital, hence two electrons per orbital.

There must therefore be as many different occupied states, which we will denote p_i, as there are particles in our system of identical Fermions, and the exchange of any two particles must change the sign of the state.

One very convenient way to do this is the *Slater determinant*, which is among the wonderfully clever computational tricks of physics that is infinitely easier to check than to come up with it oneself. With this, a Fermion state can be written as a linear combinaisons of terms of the form:

---

[1] The fact that composite particles of an even number of Fermions act like Bosons is exactly what allows some isotopes of the same element to act like superfluids and not other, as this behaviour is intimately related to the properties of Bosons, which can stack up without limit in the same quantum states.

$$|\psi\rangle_{Fermion} = \frac{1}{\sqrt{N!}} \begin{vmatrix} |\xi_1^{p_1}\rangle & |\xi_2^{p_1}\rangle & ... & |\xi_N^{p_1}\rangle \\ |\xi_1^{p_2}\rangle & |\xi_2^{p_2}\rangle & ... & |\xi_N^{p_2}\rangle \\ ... & ... & ... & ... \\ |\xi_1^{p_N}\rangle & |\xi_2^{p_N}\rangle & ... & |\xi_N^{p_N}\rangle \end{vmatrix} \tag{48}$$

We can indeed see that exchanging two particles is equivalent to exchanging two columns of the determinant, which changes its sign, and if any two $p_i = p_j$ two columns will be equal, which makes the determinant vanish thanks to its elementary linear algebra properties.

Finally, we can mention a phenomenon which we will not go too deep into: exchange interaction. This comes from the fact that the entire state of a many particle system must take on a + or - sign when it undergoes a permutation, but that the state is made up of both a spatial wavefunction and a spin state (which can be entangled). This means that if a Fermion system is in a symmetrical spin state, its wavefunction must be antisymmetrical, and the other way around. For Bosons it's the opposite, if the system is in an antisymmetrical spin state, the wavefunction must be antisymmetrical in order for the total state to be symmetrical.

As antisymmetric wavefunctions tend to be more localised than symmetric ones, they usually have higher energy. This can cause local effects like repulsion of neighbouring Fermions, the *Pauli repulsion*, even though this effect is far more general and complex when it comes to many body physics. This is the cause for both ferromagnetism and antiferromagnetism, the stability of neutron stars, and many other phenomena.

VII. BEYOND THE EMPIRICAL THEORY: A GLIMPSE AT THE DIRAC EQUATION AND THE SPIN-STATISTICS THEOREM

The addition of the new postulates that this empirical theory required was, from the start, very dissatisfactory for the physicists who developed it. Luckily for them, they wouldn't have to wait too long for more complete theories to see the light of day. Namely relativistic quantum mechanics with the Klein-Gordon and, of chief interest to us, the Dirac equation, and quantum field theory, which would become the most successful theory in history, while giving an explanation for the link between spin and the symmetrisation postulate through the spin-statistics theorem.

The first relativistic wave equation was proposed in 1926 by Klein and Gordon (even before Pauli formalised his theory of spin!), but it did not account for any additional degree of freedom for the electron.

It was derived by applying the canonical quantisation procedure directly to the relativistic equation $E^2 = p^2 + m^2$ (in the unit system where $\hbar = c = 1$), which gives:

$$\left(\partial_t^2 - \partial^i \partial_i + m^2\right) \psi = 0 \implies (\Box + m^2)\psi = 0 \tag{49}$$

Where $\Box$ stands for the D'Alembertian operator. This equation was first derived by Schrödinger, who dismissed it because of negative energy solutions which were later interpreted as antimatter. The Klein-Gordon equation accurately describes spin 0 particles, but the fact that it was second order was problematic in the eyes of Dirac. This was partly because the entirety of the information on the system was not only contained in the wavefunction, but also in its derivatives (a second order equation needs the derivative as initial condition). He wanted to get a first order equation, which meant he had to take a "square root" of the mass shell and enforce that it must be first order. To do this, he assumed that the operator acting on the wavefunction must be of the form:

$$D = \gamma_0 E - \gamma_i \, p_i - m \implies D^2 = E^2 - p^2 - m^2 \quad \text{and} \quad D\psi = 0 \tag{50}$$

This means that the objects γ_μ must obey the following properties:

$$\gamma_\mu^2 = 1 \qquad \{\gamma_\mu, \gamma_\nu\} = 0 \quad \text{for} \quad \mu \neq \nu \tag{51}$$

The smallest set of 4 objects satisfying these conditions in 3+1 spacetime are the so called Dirac matrices, which are 4x4 matrices written in terms of the Pauli matrices as:

$$\gamma^0 = \begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix} \qquad \gamma^i = \gamma^0 = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{52}$$

$$\implies (i\gamma^\mu \partial_\mu - m)\Psi = 0 \tag{53}$$

Which is the Dirac equation. The fact that these matrices are 4x4 means that the object they act on is a four component object ψ. This object is called a bispinor, or Dirac spinor, and is made up of two 2 component spinors, that behave in a similar way as those we have seen in Section V. This is remarkable, we have just found spinors in a ground up derivation of a relativistic equation. This was in fact a shock to Dirac, who did not expect that spinors, and therefore spin, would naturally arise from his derivation. As mentioned previously, it also predicted antimatter, which arose from the seemingly unphysical negative energy solutions of the equation that in fact described positive energy anti-particles.

It can be shown that these four components actually correspond to four ways objects can behave under rotation of 3+1 dimensional spacetime, linking spin to the symmetries of the universe.

When it is coupled to an electromagnetic field, gauge invariance must be enforced on the Dirac equation which undergoes the transformation $\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$, where $A = (\phi, A_i)$ is the electromagnetic four-potential. This allows us to write one equation for each of the two different spinors ψ_1 and ψ_2[1] of the Dirac spinor $\Psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$:

$$i\hbar\partial_t \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = c\begin{pmatrix}\boldsymbol{\sigma}\cdot\mathbf{\Pi}\psi_2\\ \boldsymbol{\sigma}\cdot\mathbf{\Pi}\psi_1\end{pmatrix} + q\phi\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} + mc^2\begin{pmatrix}\psi_1\\ -\psi_2\end{pmatrix} \tag{54}$$

Where $\mathbf{\Pi}$ is the minimal coupling of Section V. Next we define $\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = e^{-i\frac{mc^2}{\hbar}t}\begin{pmatrix}\psi\\ \chi\end{pmatrix}$, which we can plug into (54):

$$i\hbar\partial_t \begin{pmatrix}\psi\\ \chi\end{pmatrix} = c\begin{pmatrix}\boldsymbol{\sigma}\cdot\mathbf{\Pi}\chi\\ \boldsymbol{\sigma}\cdot\mathbf{\Pi}\psi\end{pmatrix} + q\phi\begin{pmatrix}\psi\\ \chi\end{pmatrix} - 2mc^2\begin{pmatrix}0\\ \chi\end{pmatrix} \tag{55}$$

Finally, in the non relativistic limit it can be shown that by expanding in powers of v/c the kinetic energy part is $\sim mv^2 \sim mc^2(v/c)^2$, the momentum term is $\sim pc \sim mvc \sim mc^2(v/c)$, and the electrostatic part is $\sim mc^2(v/c)^2$ generally speaking. Neglecting the terms of order $(v/c)^2$, we get for the second equation:

$$c\,\boldsymbol{\sigma}\cdot\mathbf{\Pi}\psi - 2mc^2\chi \approx 0 \implies \chi = \frac{1}{2mc}\boldsymbol{\sigma}\cdot\mathbf{\Pi}\psi \tag{56}$$

$$\implies i\hbar\partial_t\psi = \frac{(\boldsymbol{\sigma}\cdot\mathbf{\Pi})^2}{2m}\psi + q\phi\,\psi \tag{57}$$

Which is exactly what we found for the Pauli Hamiltonian! This is by no means a very rigorous derivation, but the same result can be derived formally.

We can recall that to arrive at the Pauli equation from the Schrödinger equation, we must assume that the g factor of the electron is equal to 2. Therefore the Dirac equation correctly predicts this factor of 2 since it arrives at exactly the same equation.

Adding higher order terms and taking quantum field theory into account gives the actual value for the g factor to many decimal points.

Speaking of quantum field theory, one of its achievements was to relate spin and statistics through theory. Proving it here would require the presentation of an entirely new formalism, which we cannot do without doubling the size of this paper. We will therefore simply give a qualitative discussion of the arguments used to derive it.

In quantum field theory, the fundamental object under study is a fluctuating quantum field, in which particles are oscillations. The number of particles can therefore fluctuate, and this is described by the creation and annihiation operators of the field.

First, it can be shown that if the anticommutator of the creation and annihilation operators is equal to 1, the field obeys Fermi statistics, and if it is the commutator that is equal to one it leads to Bose statistics.

Second, spinor fields which can be used to describe half integer spin particles are shown to require the aforementioned anticommutation relation in order for their energy to be positive definite, while scalar field which can be used to describe integer spin particles must obey the commutation relation above.

Therefore, spinor fields, which represent half integer spin particles, obey Fermi statistics, and scalar fields, which represent integer spin particles, obey Bose statistics. This is, in essence, one form of the spin statistics theorem.

---

[1] For which we return to the unit system where $\hbar$ and c are different from unity.

It has been shown more or less formally in different ways, but the result is always the same, half integer spins are Fermions and integer spins are Bosons.

Once again, this is not at all one of those derivations, but it simply aims at introducing the concept to the interested reader [8] [9].

LICENSING

---

[1] G. Coddens, The exact meaning of the angular-momentum and spin operators in quantum mechanics (2021), working paper or preprint, HAL open archive, hal-03323780v2.

[2] L. D. Landau and E. M. Lifshits, *Quantum Mechanics: Non-Relativistic Theory*, Course of Theoretical Physics, Vol. 3 (Butterworth-Heinemann, Oxford, 1991) isbn:978-0-7506-3539-4.

[3] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, Interscience publication, Vol. 1 (Wiley, 1977) isbn:9782705658335.

[4] R. Shankar, *Principles of quantum mechanics* (Plenum, New York, 1980).

[5] J. Jackson, *Classical Electrodynamics* (Wiley, 2021) isbn:9781119770787.

[6] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*, 3rd ed. (Cambridge University Press, Cambridge, 2018).

[7] M. Gondran and A. Gondran, A complete analysis of the Stern-Gerlach experiment using pauli spinors (2005), arXiv:quant-ph/0511276.

[8] L. Landau, E. Lifshitz, V. Berestetskii, and L. Pitaevskii, *Quantum Electrodynamics*, Course of Theoretical Physics, Vol. 4 (1994).

[9] T. Lancaster and S. J. Blundell, *Quantum Field Theory for the Gifted Amateur* (Oxford University Press, Oxford, 2014) isbn:978-0-19-969933-9.

[10] A. Martin, History of spin and statistics (2002), arXiv:hep-ph/0209068.

[11] R. Milner, A short history of spin, in *Proceedings of XVth International Workshop on Polarized Sources, Targets, and Polarimetry — PoS(PSTP2013)*, Vol. 182 (2014) p. 003.

[12] P. A. M. Dirac, An historical perspective of spin, in *Summer Study on High-energy Physics with Polarized Beams* (CERN, 1975).

[13] C. Van Vliet, *Equilibrium and Non-equilibrium Statistical Mechanics*, G - Reference, Information and Interdisciplinary Subjects Series (World Scientific Pub., 2008) isbn:9789812704771.